5-2021

Three Essays on the Performance Evaluation of Actively Managed Investment Funds

Qing Yan
University of Arkansas, Fayetteville

Three Essays on the Performance Evaluation of Actively Managed Investment Funds

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Business Administration

by

Qing Yan
South China University of Technology
Bachelor of Economics in Finance, 2010
Rensselaer Polytechnic Institute
Master of Science in Quantitative Finance and Risk Analytics, 2015

May 2021
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Alexey Malakhov, Ph.D. Dissertation Director	Timothy B. Riley, Ph.D. Dissertation Director
Wayne Y. Lee, Ph.D. Committee Member	Pu Liu, Ph.D. Committee Member

Abstract

This dissertation investigates the performance of hedge funds and actively managed U.S. equity mutual funds.

The first chapter examines the relation between hedge funds and the low beta anomaly. Different conditions in the mutual fund and hedge fund industries should lead to different approaches with respect to the low beta anomaly. I find that, unlike most mutual funds, the average hedge fund tends to benefit considerably from the anomaly. About 2.3% per year of apparent alpha for the average hedge fund can be attributed to the low beta anomaly rather than manager skill. Low skill managers are the most reliant on the anomaly to generate returns, with the most reliant underperforming the least reliant by 5.9% per year.

The second chapter uses machine learning to dynamically identify and optimally combine the predictors of hedge fund performance. The portfolio formed based on the machine learning models has an out-of-sample alpha of 7.8% per year. The importance of each predictor varies over time, but among the 22 predictors I consider, the consistently important predictors are average return, maximum return, alpha, systematic risk, and beta activity. Machine learning provides valuable, unique information about future hedge fund performance that is not captured by individual predictors.

The third chapter studies whether the quality of fund risk management can predict fund performance. I find that the risk management skills of mutual fund managers—as quantified by their funds' maximum drawdowns—are persistent and predictive of subsequent risk-adjusted performance. Funds with relatively strong past performance and relatively low past maximum drawdowns have, on average, an out-of-sample alpha of 2.68% per year. That alpha is magnified when markets are turbulent—a time during which risk management skills should be most valuable.

Investors are averse to drawdown risk. After controlling for typical measures of past performance, fund flows are still a decreasing function of maximum drawdowns, particularly among investors with greater risk aversion and during times of generally heightened risk aversion.

Acknowledgements

First, I would like to thank my advisors, Alexey Malakhov and Tim Riley, for their guidance, feedback, and support to me. Also, I would like to thank Wayne Lee and Pu Liu for their help and support to my work and academic career.

In addition, I wish to express my gratitude to Jyotishka Datta and Chanaka Edirisinghe for their help and encouragement to my research. I am also grateful to the Walton College finance seminar participants for their helpful suggestions to my work.

Moreover, I would like to thank Sean Dempsey, Xiangyu Gong, Lisa Hinrichsen, Weineng Xu, and Xin Zhang for the friendship during my Ph.D. years. Finally, my special thanks go to my parents, Qiang Yan and Yuping Zeng, for their unconditional love and everything.

Table of Contents

Chapter 1. Do Hedge Funds Bet Against Beta? 1

1.1. Introduction 1

1.2. Data and methods 8

1.2.1. Hedge fund sample 8

1.2.2. Mutual fund sample 8

1.2.3. Measuring fund performance 9

1.3. The exposures of individual hedge funds to the BAB factor 12

1.4. The performance of the average hedge fund after accounting for the BAB factor 13

1.5. The performance of the average mutual fund after accounting for the BAB factor 14

1.6. The relation between hedge funds' past performance and subsequent BAB exposure 17

1.7. The predictive power of a hedge fund's past BAB exposure 19

1.8. Hedge funds' past BAB exposures and BAB factor timing 21

1.8.1. The performance of the BAB portfolios after accounting for factor timing 22

1.8.2. Timing the BAB factor using publicly available information 24

1.9. Conclusion 27

Chapter 2. Hedge Fund Performance Prediction with Machine Learning 41

2.1. Introduction 41

2.2. Data and methods 44

2.2.1. Hedge fund data 44

2.2.2. Fund performance measure and predictors 45
2.2.3. Training and prediction 47
2.2.4. Machine learning models 47
2.2.5. Prediction evaluation 48
2.3. Prediction accuracy evaluation 49
2.4. Performance of machine learning portfolios 50
2.5. Predictor importance 53
2.5.1. Average relative importance 53
2.5.2. Time-varying importance 54
2.6. Performance of portfolios formed on individual predictors 55
2.7. Machine learning portfolios and individual portfolios 56
2.8. Conclusion 58
Chapter 3. How Risky Is Your Manager? Maximum Drawdown as Predictor of Mutual Fund Performance and Flows 75
3.1. Introduction 75
3.2. Data and methods 79
3.2.1. Mutual fund sample 79
3.2.2. Measuring maximum drawdown 80
3.2.3. Risk-adjusting performance 80
3.3. The maximum drawdowns of mutual funds 81

3.4. Can maximum drawdown predict mutual fund performance? 82

3.4.1. Returns and alphas 82

3.4.2. Subsequent MDD 83

3.4.3. Market conditions 83

3.4.4. What aspect of skill is MDD capturing? 84

3.4.5. Uniqueness of MDD 86

3.5. Do investors respond to maximum drawdown? 87

3.5.1. MDD-flow relation 87

3.5.2. MDD uniqueness 90

3.6. Conclusion 90

References 102

List of Tables

Table 1.1: The performance of an equal weight portfolio of all hedge funds 28

Table 1.2: The performance of an equal weight portfolio of all ‘traditional’ mutual funds......... 29

Table 1.3: The performance of an equal weight portfolio of all ‘alternative’ mutual funds 30

Table 1.4: The performance of the top and bottom alpha decile portfolios................................. 31

Table 1.5: The characteristics of the top and bottom BAB hedge funds 32

Table 1.6: The performance of the BAB portfolios using LASSO... 33

Table 1.7: The performance of the BAB portfolios using OLS.. 34

Table 1.8: The performance of the BAB portfolios using LASSO after accounting for market timing 35

Table 1.9: The performance of the BAB portfolios using OLS after accounting for market timing 36

Table 1.10: The TED spread and the BAB exposures of the BAB portfolios 37

Table 2.1: Summary statistics: 1994 – 2018.. 60

Table 2.2: Prediction accuracy... 62

Table 2.3: Comparing predictions using Diebold-Mariano tests 63

Table 2.4: Performance of machine learning portfolios 64

Table 2.5: Performance of machine learning portfolios – subperiods 65

Table 2.6: Performance of portfolios formed on individual predictors – FH8 alpha 66

Table 2.7: Performance of portfolios formed on individual predictors – other measures 67

Table 2.8: Machine learning portfolio (LASSO) and classical portfolio...................................... 69

Table 2.9: Machine learning portfolio (OLS) and classical portfolio... 70

Table 3.1: Characteristics of mutual funds with low and high MDD... 92

Table 3.2: Alphas of portfolios formed based on past performance and MDD............................ 93

Table 3.3: MDDs of portfolios formed based on past performance and MDD 94

Table 3.4: Performance as a function of market conditions 95

Table 3.5: What aspects of skill is MDD capturing? 96

Table 3.6: Comparison of MDD and other measures as performance predictors......................... 97

Table 3.7: MDD-flow relation 98

Table 3.8: Relation between flows, MDD, and MDD-related measures 99

List of Figures

Figure 1.1: The frequency of the BAB and FH8 factors being selected by LASSO regressions . 38

Figure 1.2: The distribution of hedge funds' BAB exposures 39

Figure 1.3: The statistical significance of hedge funds' BAB exposures 40

Figure 2.1: Value of machine learning portfolios 71

Figure 2.2: Predictor importance 72

Figure 2.3: Predictor time vary importance – relative importance 73

Figure 2.4: Predictor time vary importance – absolute importance 74

Figure 3.1: Average unadjusted maximum drawdown of mutual funds 100

Figure 3.2: Cumulative return on portfolios formed based on past performance and MDD 101

Chapter 1. Do Hedge Funds Bet Against Beta?

Alexey Malakhov, Timothy B. Riley, Qing Yan

1.1. Introduction

The capital asset pricing model (CAPM) assumes all investors can use leverage to achieve their optimal portfolio given their risk preferences, but many investors face constraints on leverage. Frazzini and Pedersen (2014) contend that leverage constrained investors will overweight high beta assets in their portfolios, which will cause high beta assets to be overpriced relative to low beta assets. To capture this low beta anomaly—which was first shown empirically for U.S. equities in the early 1970s—Frazzini and Pedersen (2014) construct pricing factors that long a leveraged portfolio of low beta assets and short a deleveraged portfolio of high beta assets.[1] While there is debate about the actual mechanism that causes the low beta anomaly, the betting-against-beta (BAB) factors still generate economically and statistically significant positive average returns across many different asset classes and countries.[2]

The relation between mutual funds and the low beta anomaly has received extensive study. Baker, Bradley, and Wurgler (2011) argue that the combination of a leverage constraint and a mandate to beat a fixed benchmark will discourage an institutional investor from capitalizing on the anomaly. Christoffersen and Simutin (2017) find that domestic equity mutual funds, which are subject to leverage constraints and benchmark mandates, tend to tilt their portfolios toward high beta stocks to increase the probability of beating their benchmark. Consistent with that result, both Karceski (2002) and Frazzini and Pedersen (2014) show evidence that domestic equity mutual funds tend to hold stocks with betas greater than one. Similarly, Choi and Kronlund (2018) find

[1] For early evidence of the low beta anomaly, see, for example, Friend and Blume (1970); Black, Jensen, and Scholes (1972); and Haugen and Heins (1972, 1975).
[2] Cederburg and O'Doherty (2016); Hong and Sraer (2016); Bali, Brown, Murray, and Tang (2017); Liu, Stambaugh, and Yuan (2018); and Han (2019), among others, propose alternative explanations for the low beta anomaly.

that corporate bond mutual funds tend to tilt their portfolios toward higher yielding bonds. While this tendency should not generate true alpha, it is rational from the perspective of mutual fund managers. Sensoy (2009) shows that a mutual fund's performance relative to its fixed benchmark has a large effect on net asset flows, and Lee, Trzcinka, and Venkatesan (2019) find that same performance is frequently linked to a mutual fund manager's compensation.

In this paper, we consider the relation between hedge funds and the low beta anomaly. The conditions of the hedge fund industry are significantly different from those of the mutual fund industry. First, unlike mutual funds, hedge funds have the tools necessary to take advantage of the low beta anomaly. Hedge funds are less leverage constrained than mutual funds and less restricted with regard to short selling. Ang, Gorovyy, and van Inwegen (2011) show that the average hedge fund uses substantial leverage, while mutual funds are subject to strict regulatory leverage constraints.[3] Consequently, hedge funds are able—if they choose—to capitalize on the low beta anomaly by leveraging positions that are long low beta assets and short high beta assets.

Second, unlike mutual funds, hedge funds have managers with greater incentive to take advantage of the low beta anomaly. Rather than the fixed benchmark driven compensation common to mutual funds managers (Ma, Tang, and Gomez, 2019, and DeWoody, 2019), hedge fund managers are typically compensated using a 2/20 fee structure, conditional on high water marks and hurdle rates. Moreover, Agarwal, Green, and Ren (2018) show that investors making capital allocations often treat hedge fund returns attributable to non-traditional factor exposures as alpha.[4] Cochrane (2011, pg. 1087) illustrates this point as follows:

> *I tried telling a hedge fund manager, "You don't have alpha. Your returns can be replicated with a value-growth, momentum, currency and term carry, and short-vol strategy." He*

[3] The use of leverage by mutual funds is limited under Section 18 of the Investment Company Act of 1940.
[4] In this respect, mutual fund and hedge fund investors are similar as mutual fund investors also tend to treat some betas as alpha (e.g., Berk and van Binsbergen, 2016; Barber, Huang, and Odean, 2016; and Ben-David, Li, Rossi, and Song, 2019).

> *said, "'Exotic beta' is my alpha. I understand those systematic factors and know how to trade them. My clients don't."*

Considering their personal compensation structure and their investors' behavior, hedge fund managers should have a strong incentive to focus on generating absolute returns, rather than the relative returns that are the focus for mutual fund managers. Since strategies based on the low beta anomaly can be expected to generate positive absolute returns, hedge fund managers should choose—if able—to capitalize on the anomaly. Given both their ability and their incentive to capitalize, we hypothesize that hedge funds will tend to "bet against beta."[5]

We begin our analysis of that hypothesis by examining whether a broad 'All Assets' BAB factor is relevant for hedge funds. Specifically, we add that BAB factor to the Fung and Hsieh (2004) eight-factor hedge fund model and use LASSO (least absolute shrinkage and selection operator) regressions to select the relevant factors for each hedge fund in our sample.[6] We find that the BAB factor has a higher probability of being selected as relevant than all but one of the Fung and Hsieh (2004) factors (the emerging markets factor). When the BAB factor is selected, the hedge fund's exposure to the factor is positive about 94% of the time. That is, consistent with our hypothesis, hedge funds do tend to "bet against beta."

Accounting for the BAB exposures of hedge funds has a noticeable effect on measures of performance. The alpha of an equal weight portfolio of all hedge funds decreases by 2.3% per year after adding the BAB factor to the Fung and Hsieh (2004) model. As of the first quarter of 2018, the hedge fund industry managed about $3.2 trillion.[7] Therefore, about $73 billion dollars in

[5] Arguments based on the leverage aspect of our theory have also been put forth to explain why other investors tend to "bet against beta." For example, Frazzini and Pedersen (2014) show that private equity funds tend to have a beta less than one, and Frazzini, Kabiller, and Pedersen (2018) show a large, positive BAB exposure for Berkshire Hathaway.
[6] LASSO regression is introduced in Tibshirani (1996) and Efron, Hastie, Johnstone, and Tibshirani (2004).
[7] Data from Hedge Fund Research's *Global Hedge Fund Industry Report* https://www.hedgefundresearch.com/news/hfr-global-hedge-fund-industry-report-q1-2018-published

apparent alpha generated by hedge funds each year can be attributed to positive BAB exposure. While Ibbotson, Chen, and Zhu (2011) suggest an alpha of 3.0% per year for the average hedge fund, we find a statistically insignificant post-BAB average alpha of 1.6% per year. Our results here match prior work showing that the alpha of the average hedge fund is often overstated (for instance, Malkiel and Saha, 2005; Aragon, 2007; Fung, Hsieh, Naik, and Ramadorai, 2008; and Aiken, Clifford, and Ellis, 2013). While we are not the first to document that accounting for BAB exposure decreases the average hedge fund alpha (see Joenvaara, Kauppila, Kosowski, and Tolonen, 2019), we are the first to build and empirically investigate a conceptual framework explaining that result.[8]

Along those lines, the results for hedge funds stand in contrast to those for actively managed mutual funds with 'traditional' strategies, such as domestic equity and fixed income. We find that the alpha of an equal weight portfolio of all 'traditional' mutual funds is unaffected by adding the BAB factor to the Fung and Hsieh (2004) model. LASSO regressions do not even select the BAB factor as relevant, which implies that the average 'traditional' mutual fund derives neither significant costs nor benefits from the low beta anomaly. However, among the relatively small number of mutual funds following hedge-fund-like 'alternative' strategies, such as absolute return and market neutral, the results are similar to those for hedge funds—albeit economically and statistically weaker. An equal weight portfolio of all 'alternative' mutual funds has a large, positive BAB exposure, which when accounted for decreases the portfolio's alpha by about 1.4% per year. Considering that 'alternative' mutual funds often resemble hedge funds in benchmark (e.g., a positive return) and investment restrictions (e.g., more use of leverage, derivatives, and shorting),

[8] We provide more detail on the differences between our paper and Joenvaara, Kauppila, Kosowski, and Tolonen (2019) in Section 4.

that result is consistent with the idea that weakening a combination of barriers will allow an institutional investor to capitalize on the low beta anomaly.[9]

We further hypothesize that the tendency to "bet against beta" will be strongest among hedge fund managers with relatively low skill. Given that hedge fund investors often treat returns generated through exotic factor exposures as alpha, relatively low skill hedge fund managers should have the incentive to generate returns through BAB exposure to compensate for not being able to generate true alpha.[10] That behavior would be aligned with the findings of Titman and Tiu (2011)—who show that relatively low skill hedge fund managers are more reliant on factor exposures to generate returns—and Sun, Wang, and Zheng (2012)—who find the same low skill managers also choose to pursue less unique strategies. It would, however, run counter to the positive relation between hedge fund systematic risk and alpha shown by Bali, Brown, and Caglayan (2012).

Because the true skill of a hedge fund manager is unobservable, we begin our analysis of this hypothesis by using a hedge fund's past performance as a proxy for skill. We find a large and highly statistically significant exposure to the BAB factor in the bottom decile of funds sorted by past performance, while the BAB exposure of funds in the top decile is much smaller and not statistically significant. The alpha of an equal weight portfolio of worst performers decreases by a statistically significant 3.7% per year after adding the BAB factor to the Fung and Hsieh (2004) model, while the decrease in alpha for an equal weight portfolio of best performers is a statistically insignificant 1.3% per year. Hence, consistent with our hypothesis, the worst performing (and

[9] While 'alternative' mutual funds must still comply with all U.S. Securities and Exchange Commission (SEC) rules, they typically have fewer institutional restrictions, which can often be numerous and binding (see, for example, Almazan, Brown, Carlson, and Chapman, 2004).

[10] Berk and van Binsbergen (2015) argue that, in equilibrium, alpha is unrelated to skill, but (i) their model is subject to significant debate—particularly with respect to diseconomies of scale (e.g., Phillips, Pukthuanthong, and Rau, 2018, and Adams, Hayunga, and Mansi, 2018)—and (ii) their model is specifically focused on the mutual fund industry. In our analysis, we operate from the traditional viewpoint that there is a positive relation between alpha and skill.

presumably the lowest skill) hedge fund managers rely the most on the low beta anomaly to generate returns.

The arguments above also suggest that BAB exposure itself should signal skill: a large BAB exposure can indicate an attempt by a relatively low skill hedge fund manager to substitute for true alpha. Therefore, we further our analysis by considering the predictive power of a hedge fund's BAB exposure. Jordan and Riley (2015) find that, among domestic equity mutual funds, those that invest in relatively low risk stocks outperform those that invest in relatively high risk stocks if standard evaluation models are used. However, they also show that the apparent difference in performance is eliminated if the models are modified to account for the low beta anomaly. In contrast, we find in a similar test that accounting for the low beta anomaly does not eliminate, but rather reveals, hedge fund alpha.

Using the Fung and Hsieh (2004) model, an equal weight portfolio of hedge funds in the lowest decile of past BAB exposure does not outperform an equal weight portfolio of hedge funds in the highest decile. But after the BAB factor is added to that model, the lowest decile portfolio outperforms the highest decile portfolio by a substantial margin of 5.9% per year. To understand the economic size of that result, note that the S&P 500 had an average total return of 5.8% per year over the same time period (January 1998 through March 2012). The large difference in performance that appears between the two portfolios when BAB exposure is not treated as alpha supports the idea that a hedge fund's BAB exposure can signal managerial skill, with high BAB exposures prevalent among managers with relatively low skill.

Subsequent analyses using models styled after Henriksson and Merton (1981) show that hedge funds in the highest decile of past BAB exposure successfully time the BAB factor. Nevertheless, even after accounting for the gains from BAB factor timing, the performance of the highest decile portfolio is still poor compared to the lowest decile portfolio. Despite not generating any returns from timing the BAB factor, the lowest decile portfolio still has an average total return from active management—the sum of the returns from all timing and non-timing related active management decisions—that is 4.9% per year greater than that of the highest decile portfolio. More generally, the lowest decile portfolio delivers a greater average excess return and Sharpe ratio compared to the highest decile portfolio.[11] All of these performance differences are consistent with the prior evidence suggesting that high BAB exposure is a signal of relatively low hedge fund manager skill.

Further supporting that claim, we find that a substantial portion of hedge funds' BAB factor timing is achieved using public information, rather than unique managerial insight. Frazzini and Pedersen (2014) document a negative relation between the lagged level of the TED spread—a proxy for funding constraint tightness—and the return on the BAB factor. While hedge funds in the lowest past BAB exposure decile portfolio show some evidence of modifying their BAB exposure based on the lagged TED spread, we observe that the hedge funds in the highest decile portfolio make relatively large modifications consistent with public-information-based timing. When the lagged TED spread is relatively high, the BAB exposure of the highest decile portfolio is 0.59, but when the lagged TED spread is relatively low, the exposure is 1.30. If we accept Ferson and Schadt's (1996, pg. 428) argument that returns from public-information-based timing should not be treated as alpha "because investors can replicate on their own any strategy which depends on public information," then the lowest decile portfolio outperforms the highest decile portfolio by over 6% per year, regardless of the lagged TED spread. That result adds additional support to the idea that relatively low skill hedge fund managers tend to use a large amount of BAB exposure.

[11] The highest past BAB exposure decile portfolio has an annualized average excess return of 4.3% and a Sharpe ratio of 0.45. The same values for the lowest past BAB exposure decile portfolio are 6.3% and 0.64.

As a whole, it is clear that hedge funds "bet against beta." Many hedge funds favor low beta assets over high beta assets, while almost no hedge funds show the opposite preference. As a result, the performance of the average hedge fund is significantly overstated by models that fail to account for the low beta anomaly. Reliance on the anomaly to generate returns is most prevalent among relatively low skill hedge fund managers, which suggests that much of those hedge funds' already low apparent alpha is not alpha at all.

1.2. Data and methods

1.2.1. Hedge fund sample

Following Duanmu, Malakhov, and McCumber (2018), we use Bloomberg to collect extensive hedge fund data from 1994 to 2015. The data includes monthly net returns and common characteristics, such as assets under management and style.[12] Like other databases, hedge funds report data to Bloomberg on voluntary basis, although Bloomberg requires that all reporting funds provide performance since inception. To reduce backfill bias, we drop the first 24 months of each fund's returns.[13] We include both live and defunct funds in our sample to control for survivorship bias, and we remove all funds of funds. The BAB factor that we use in all of our tests is only available through March 2012, so our sample period is January 1996 through March 2012. The final sample contains 10,095 unique funds and 456,861 fund-month observations.

1.2.2. Mutual fund sample

We use the Center for Research in Security Prices (CRSP) Survivor-Bias-Free U.S. Mutual Fund database to build our sample of actively managed mutual funds. We exclude any fund that CRSP identifies as an index fund, ETF, or variable annuity; drop all funds of funds and money

---

[12] Hedge fund returns are calculated from fund net asset values (NAVs) converted into U.S. dollars.
[13] The date upon which a hedge fund first reports data to Bloomberg is unavailable. Therefore, we do not know the exact number of backfilled returns for any given fund. The 24-month correction is consistent with both Jagannathan, Malakhov, and Novikov (2010) and Titman and Tiu (2011).

market funds; and only include a fund in our sample once it is at least two years old and has at some point reached at least \$20 million in assets (to address the incubation bias identified by Evans, 2010). Multiple share classes of a mutual fund are collapsed into a single fund with (i) fund assets being the sum of the assets across all share classes and (ii) all other fund characteristics, including net return, being asset-weighted averages of the share class values.[14] We match the time period for our mutual fund sample to that of the hedge fund sample—January 1996 through March 2012—which results in it having 8,387 unique funds and 831,772 fund-month observations.

We subdivide the full sample of actively managed mutual funds into ‘alternative’ and ‘traditional’ groups based on the fund’s investment strategy. The funds in the ‘alternative’ group follow hedge-fund-like strategies and are identified based on Lipper classification.[15] All funds not identified as ‘alternative’ are placed in the ‘traditional’ group. Funds in the ‘traditional’ group can be broadly classified using CRSP objective codes as either domestic equity (43% of the ‘traditional’ sample), foreign equity (13%), fixed income (34%), or mixed/other (10%). The ‘alternative’ group has 223 unique funds and 14,152 fund-month observations, and the ‘traditional’ group has 8,164 unique funds and 817,620 fund-month observations.

1.2.3. Measuring fund performance

We use the Fung and Hsieh (2004) eight-factor hedge fund model (henceforth, the FH8 model) as the baseline for evaluating fund performance:

$$\begin{aligned} r - r_f = \alpha + \beta_1 SP500 + \beta_2 SizeSpread + \beta_3 EmergMkt + \beta_4 10Year \\ + \beta_5 CreditSpread + \beta_6 BondTrend + \beta_7 FxTrend + \beta_8 ComTrend \\ + \varepsilon \end{aligned} \tag{1.1}$$

[14] The net returns we consider in our analysis of mutual funds do not reflect any effects from front- or back-end loads.
[15] Specifically, a mutual fund is considered ‘alternative’ if it uses any of the following Lipper classifications: ABR, ACF, AED, AGM, ALT, AMS, BBBL, CMS, DL, DSB, EMN, LSE, MFF, SESE, or SFI.

r is the monthly net return of a given fund, r_f is the risk-free rate (proxied by the return on 30-day US Treasury bills), and α is the fund's alpha. The included factors are:

(1) $SP500$, the S&P 500 index return minus the risk-free rate;

(2) $SizeSpread$, the Russell 2000 index return minus the S&P 500 index return;

(3) $EmergMkt$, the MSCI Emerging Market index return minus the risk-free rate;[16]

(4) $10Year$, the 10-year U.S. Treasury bond portfolio return minus the risk-free rate;

(5) $CreditSpread$, the Citi BBB corporate bond index return minus the Fama U.S. Treasury bond portfolio return;

(6) $BondTrend$, $FxTrend$, and $ComTrend$, three trend-following factors constructed from lookback straddles on futures contracts on bonds, currencies, and commodities.[17]

All the data on returns is from CRSP and Bloomberg, while the three trend-following factors are courtesy of David Hsieh.[18]

To evaluate the relation between funds and the low beta anomaly, we modify that baseline model by adding the Frazzini and Pedersen (2014) betting-against-beta factor (henceforth, the FH8+BAB model):

$$\begin{aligned} r - r_f = \alpha + \beta_1 SP500 + \beta_2 SizeSpread + \beta_3 EmergMkt + \beta_4 10Year \\ + \beta_5 CreditSpread \\ + \beta_6 BondTrend + \beta_7 FxTrend + \beta_8 ComTrend + \beta_9 BAB + \varepsilon \end{aligned} \tag{1.2}$$

BAB is the monthly return on a portfolio constructed by longing a leveraged portfolio of low beta assets and shorting a deleveraged portfolio of high beta assets.

---

[16] *EmergMkt* was not added into the model until Edelman, Fung, Hsieh, and Naik (2012). As shown in our later results, *EmergMkt* is an important factor in explaining the returns of the average hedge fund.
[17] See Fung and Hsieh (2001) for more information on the trend-following factors.
[18] We thank David Hsieh for making the trend-following factor data available at his website. https://faculty.fuqua.duke.edu/~dah7/HFRFData.htm

Frazzini and Pedersen (2014) construct a variety of BAB factors using different assets classes and markets. Given the broad diversification of hedge funds, we use their ‘All Assets’ BAB factor in our analysis.[19] Prior work—focusing on the ‘U.S. Equity’ BAB factor rather than the ‘All Assets’ BAB factor—has disagreed on whether the BAB factor is investable. For example, Auer and Schuhmacher (2015, pg. 30) claim a BAB-based strategy can be successfully implemented using only “the 30 highly liquid stocks of the Dow Jones Industrial Average,” while Novy-Marx and Velikov (2016, 2019) claim that a successful BAB-based strategy cannot be implemented because of a necessary overreliance on small-cap and micro-cap stocks.[20] By examining net-of-fee returns experienced by actual investors rather than hypothetical portfolios, we can examine whether generating significant BAB exposure is possible, regardless of whether the exact BAB factor can be inexpensively purchased.

Throughout the paper, we estimate Eq. (1.1) and Eq. (1.2) using standard OLS regressions. We also estimate the models using a procedure designed to eliminate irrelevant factors, since not all of the factors in the models are relevant for all hedge funds at all points in time. To identify the relevant factor exposures for a given individual hedge fund or a given portfolio of hedge funds, we follow Li and Malakhov (2017) and first use a LASSO regression as a factor selection method. We then use an OLS regression to estimate the exposures to the factors selected by the LASSO regression. We refer to the entirety of this two-step process as a LASSO regression to differentiate it from the standard OLS regression.

[19] We thank the authors for making the returns on the ‘All Assets’ BAB factor available through March 2012 on Applied Quantitative Research’s (AQR’s) website. Extending forward the time-series of the factor’s returns ourselves is cost prohibitive. See Section 3 and Table 8 of Frazzini and Pedersen (2014) for details on the factor’s construction. https://www.aqr.com/Insights/Datasets/Betting-Against-Beta-Original-Paper-Data

[20] Hou, Xue, and Zhang (2020) likewise show that the returns on the BAB factor are sensitive to how the formation procedure handles stocks with low market capitalization.

1.3. The exposures of individual hedge funds to the BAB factor

Our examination of whether the BAB factor is an important factor for hedge funds begins by using LASSO regressions to select the relevant factors for each fund. Specifically, we start by requiring that each fund have at least 24 monthly return observations, which leaves us with 6,163 unique funds. We then run a LASSO regression on each fund's full return history using the FH8+BAB model. Figure 1.1 shows how often each of the nine factors in the model is selected by the LASSO regressions and whether the resulting factor exposure is positive or negative. Indicating the factor's importance in explaining hedge fund returns, the BAB factor is selected for 36% of hedge funds, which is more frequent than all other factors, except for the emerging market factor. Indicating whether hedge funds bet for or against beta, among the funds for which the BAB factor is selected, about 94% have a positive exposure (i.e., when betting, hedge funds tend to "bet against beta").

Next, we assess the distribution of BAB exposure within the hedge fund industry. To generate that distribution, we run OLS regressions on the same set of funds using the same model. Figure 1.2 shows that there is significant variation in BAB exposure across funds, but that most funds have a positive exposure. After trimming the distribution at the 1% level, BAB exposure varies between −2.0 and 3.2; however, about 52% of funds have an exposure between 0.0 and 1.0 and another 24% have an exposure greater than 1.0. Untabulated results show that BAB exposure has a mean of 0.51, a median of 0.39, and a standard deviation of 0.80.

We evaluate the statistical significance of the above BAB exposures in Figure 1.3. Among the 76% of hedge funds with a positive exposure to the BAB factor, about 49% are statistically significant at the 10% level. Put another way, about 37% (=0.76 x 0.49) of hedge funds have a statistically significant positive BAB exposure. In comparison, only 3% of hedge funds have a statistically significant negative BAB exposure. Like in Figure 1.1, if the results in Figures 1.2 and 1.3 are considered in combination, it is clear that the BAB factor is an important factor for hedge funds and that hedge funds typically have a positive BAB exposure.

1.4. The performance of the average hedge fund after accounting for the BAB factor

Given the relevance of the BAB factor in explaining hedge funds' returns, it is reasonable to expect that accounting for BAB exposure would meaningfully impact hedge funds' alphas. To test the effect on performance of accounting for the BAB factor, we create an equal weight portfolio of all hedge funds and measure its alpha using both the FH8 model and the FH8+BAB model. Table 1.1 reports the results from this analysis using both LASSO regressions and OLS regressions. Since, in this instance, the results are very similar across those methods, we only discuss the LASSO regression results.

The BAB factor is selected by the LASSO regression as relevant for the portfolio, and the portfolio has a large, positive exposure to the factor of 0.40 (*t*-stat = 6.10). The magnitude of the BAB factor's exposure is more than double that of any other selected factor. Adding the BAB factor to the model decreases the alpha of the portfolio from 0.30% per month (*t*-stat = 3.71) to 0.12% per month (*t*-stat = 1.43). That difference amounts to a decrease in alpha of 2.3% per year (*t*-stat = 2.32). Total hedge fund industry assets were about $3.2 trillion in the first quarter of 2018, which suggests that about $73 billion per year in apparent alpha can be attributed to hedge funds' positive exposure to the BAB factor.[21] The statistically insignificant alpha of about 1.4% per year that remains for the average hedge fund after accounting for the BAB factor may still overstate performance as the FH8 model tends to produce upward biased alphas (Bhardwaj, Gorton, and Rouwenhorst, 2014).

[21] Portfolios of hedge funds with above and below median assets under management both have large, positive average BAB exposures.

Our results to this point run counter to Blitz's (2018) claim that hedge funds actually bet on, rather than against, beta.[22] Our results, however, are consistent with Joenvaara, Kauppila, Kosowski, and Tolonen (2019), who find that the average hedge fund has a positive exposure to a 'Global Equity' BAB factor.[23] While similar in limited respects, we note that the overall goal and findings of our paper are very different from Joenvaara, Kauppila, Kosowski, and Tolonen (2019). Our focus is the relation between hedge funds and the low beta anomaly, while they briefly evaluate the BAB factor as part of an analysis focused on the sensitivity of results to hedge fund database choices. Regarding the potential overlap, our study differs from theirs in both execution and results. In terms of execution, when testing the BAB factor, we use a different base factor model (augmented global Carhart, 1997, vs. Fung and Hsieh, 2004) and a much broader BAB factor ('Global Equity' vs. 'All Assets').[24,25] In terms of results, we find both a larger average BAB exposure (0.11 vs. 0.40) and a larger impact on alpha (1.3% per year vs. 2.3% per year).

1.5. The performance of the average mutual fund after accounting for the BAB factor

We previously argued that, unlike mutual funds, hedge funds have both the ability and incentive to capitalize on the low beta anomaly. Here, we evaluate the extent to which ability and incentive matter by replicating the analysis from the prior section using two distinct groups of

[22] The results in Blitz (2018) do not hold in our sample. We find in untabulated results that including our replication of his 'Low Volatility – High Volatility' factor in the Fung and Hsieh (2004) model has no effect on the alpha of an equal weight portfolio of all hedge funds. The factor loading itself is not statistically significant.

[23] More generally, our results are consistent with prior work showing that hedge funds are able to identify and take advantage of mispricing. See, for example, Brunnermeier and Nagel (2004); Kokkonen and Suominen (2015); Sias, Turtle, and Zykaj (2016); Caglayan, Celiker, and Sonaer (2018); Cao, Chen, Goetzmann, and Liang (2018); and Cao, Liang, Lo, and Petrasek (2018).

[24] If we reperform our previous analysis using the global Carhart (1997) model; the augmented global Carhart (1997) model (excluding the 'Global Equity' BAB factor); the U.S. Carhart (1997) model; or the Cremers, Petajisto, and Zitzewitz (2013) seven-factor model as our baseline, we still find that hedge funds have a positive exposure to the 'All Assets' BAB factor that when accounted for substantially decreases alpha. We also still find the same trend in BAB exposure that we discuss in the next section.

[25] Untabulated analysis that includes both BAB factors in the FH8 model shows that the average hedge fund has a larger exposure to the 'All Assets' BAB factor (0.08 vs. 0.26). Placing just the 'Global Equity' BAB factor into the FH8 model leaves the alpha of the average fund positive and statistically significant.

mutual funds: those with ‘traditional’ strategies and those with ‘alternative’ strategies. Because of the conditions under which ‘traditional’ mutual funds operate, we do not expect them to take advantage of the low beta anomaly through positive exposure to the BAB factor. However, among the hedge-fund-like ‘alternative’ mutual funds, who are subject to fewer investment constraints and pursue absolute returns, we expect a tendency toward positive BAB exposure (although that tendency should still be weaker than that of the even less constrained hedge funds).[26]

We first form an equal weight portfolio of all ‘traditional’ mutual funds and evaluate performance with and without considering exposure to the BAB factor. To maintain a clean comparison with our previous hedge fund results, we use the FH8 model as our baseline and use the same time period.[27] Table 1.2 reports the results from both LASSO regressions and OLS regressions. Consistent with our expectations, the BAB factor is of limited importance in explaining the performance of the average ‘traditional’ mutual fund. The BAB factor is not selected as relevant by the LASSO regression, and while the OLS regression does show a positive and statistically significant BAB exposure of 0.08 (*t*-stat = 2.25), the impact on alpha of accounting for that exposure is economically trivial and statistically insignificant (0.03% per month, *t*-stat = 0.83).[28,29] In addition, adding the BAB factor to the FH8 model when using the OLS regressions has no impact on the adjusted R^2.

---

[26] See, for example, Agarwal, Boyson, and Naik (2009); Clifford, Jordan, and Riley (2013); and Huang and Wang (2013) for more discussion on mutual funds with ‘alternative’ strategies.

[27] The results are similar (i) if we use only domestic equity mutual funds and (ii) if we use only domestic equity mutual funds and replace the ‘All Assets’ BAB factor with the ‘U.S. Equity’ BAB factor.

[28] When the LASSO regressions select the same factors from both the FH8 set and the FH8+BAB set, we do not report differences between those results (because there are, of course, no differences).

[29] Using the augmented global Carhart (1997) hedge fund model of Joenvaara, Kauppila, Kosowski, and Tolonen (2019) (excluding the ‘Global Equity’ BAB factor), the exposure to the ‘All Assets’ BAB factor when using the OLS regressions is −0.11 (*t*-stat = −2.42).

In Table 1.3, we repeat the above tests using our sample of 'alternative' mutual funds.[30] Because the results from the LASSO regressions and the OLS regressions are very similar in this instance, we only discuss the results from the LASSO regressions. In line with our expectations, the BAB factor is selected by the LASSO regression as relevant for the equal weight portfolio of 'alternative' mutual funds. The BAB exposure of the portfolio is 0.25 (*t*-stat = 7.31). After adding the BAB factor to the baseline FH8 model, the already negative alpha of –0.08% per month (*t*-stat = –1.81) decreases to –0.20% per month (*t*-stat = –4.66). That decrease in alpha of –0.12% per month is both statistically (*t*-stat = –2.76) and economically significant (–1.4% per year). Like hedge funds, 'alternative' mutual funds extensively rely on BAB exposure to generate returns, albeit to a lesser degree.

In the big picture, there is a notable trend in the increasing level of BAB exposure when moving from 'traditional' mutual funds to 'alternative' mutual funds to hedge funds. In untabulated results, we find (i) that 'alternative' mutual funds have a BAB exposure that is 0.16 (*t*-stat = 4.25) greater than that of 'traditional' mutual funds and (ii) that hedge funds have a BAB exposure that is 0.16 (*t*-stat = 2.62) greater than that of 'alternative' mutual funds. The availability of the necessary tools to capitalize and the incentive to capitalize on the low beta anomaly is also trending in the same direction. Those trends matching is consistent with our hypothesis on why mutual funds and hedge funds should have different approaches to the low beta anomaly. 'Alternative' mutual funds occupy the middle ground between 'traditional' mutual funds and hedge funds with respect to their benchmarks and investment constraints, and consequently, they also occupy the middle ground with respect to BAB exposure.

[30] Over the sample period, there is a significant increase in the number of 'alternative' mutual funds. There are 23 such funds when the sample begins and 176 when the sample ends. If the regressions are weighted by the number of funds in the sample in a given month, the results are similar to those presented.

1.6. The relation between hedge funds' past performance and subsequent BAB exposure

The inclusion of the BAB factor into the FH8 model causes a sizeable decrease in the alpha of the average hedge fund. We consider here whether the impact of the BAB factor on alpha varies depending on a hedge fund's past performance. In regard to funds that have outperformed in the past, Kosowski, Naik, and Teo (2007) and Jagannathan, Malakhov, and Novikov (2010) both find outperformance continues in the future. It is possible that the persistent performance of funds with strong past performance is driven by their exposure to the BAB factor, rather than manager skill. In regard to funds that have underperformed in the past, Titman and Tiu (2011) show relatively low skill hedge fund managers are more reliant on factor exposures to generate returns. Therefore, it is also possible that funds with poor past performance are heavily reliant on BAB exposure, rather than manager skill.

We form two portfolios to test how accounting for the BAB factor affects the subsequent performance of hedge funds with past outperformance or underperformance. The 'Top α' portfolio is an equal weight portfolio of all hedge funds in the highest decile of past performance, and the 'Btm α' portfolio is an equal weight portfolio of all hedge funds in the lowest decile. To measure past performance, we estimate an alpha for each fund at the beginning of each month using LASSO regression, the past 24 monthly returns, and the FH8+BAB model. These portfolios are not formed in the first two years of our sample period, since 24 months of past returns are not available during that time.

Table 1.4 reports the alphas and factor exposures for these portfolios using both the FH8 model and the FH8+BAB model. As before, LASSO regressions and OLS regressions are both considered. Using LASSO regression, the BAB factor is not selected for the 'Top α' portfolio. Accordingly, the economically large and statistically significant positive alpha of the 'Top α' portfolio is unchanged by the addition of the BAB factor. LASSO regression does select the BAB factor for the 'Btm α' portfolio. That portfolio's exposure to the BAB factor is 0.80 (*t*-stat = 5.89)—more than three times the magnitude of the 'Btm α' portfolio's next largest exposure. As a result of that substantial BAB exposure, adding the BAB factor to the FH8 model decreases the alpha of the 'Btm α' portfolio from economically and statistically zero to −0.27% per month (*t*-stat = −1.53).[31]

Turning to the OLS regressions, the 'Top α' portfolio does have a positive exposure to BAB factor of 0.25 (*t*-stat = 1.52), but accounting for that exposure has a limited impact on conclusions about that portfolio's performance. Alpha for the 'Top α' portfolio is 0.74% per month (*t*-stat = 3.70) using the FH8 model and 0.63% per month (*t*-stat = 3.01) using the FH8+BAB model. The difference in alpha of 0.11% per month between those two models is not statistically significant (*t*-stat = 0.50), and in both cases, performance still persists. Results for the 'Btm α' portfolio using OLS regressions are similar to those from using LASSO regressions. The BAB exposure of the 'Btm α' portfolio is large and positive, and accounting for that exposure decreases that portfolio's alpha by 0.31% per month (*t*-stat = 1.73).

The above results suggest that (i) the persistent performance of the best performing funds is not the result of BAB exposure and (ii) the worst performing funds do heavily rely on BAB exposure.[32] To test the robustness of these findings, we repeated the analysis using estimates of past performance from OLS regressions rather than LASSO regressions and using estimates of past performance from the FH8 model rather than the FH8+BAB model. Results from those

[31] When using the LASSO regressions, we cannot always test the statistical significance of the changes from adding the BAB factor to the FH8 model. Such a test is possible when the LASSO regressions only add factors in the FH8+BAB model that were not selected in the FH8 model, but such a test is not possible when factors are simultaneously added and subtracted.

[32] These results do not necessarily imply an investable strategy, as Joenvaara, Kosowski, and Tolonen (2019) find that constraints (e.g., rebalancing frequency, minimum investment, closed-to-new-investments) keep investors from easily exploiting this type of hypothetical portfolio.

alternative specifications are consistent with those presented in Table 4. Furthermore, if we repeat this analysis separately for each hedge fund style (as classified by Agarwal, Daniel, and Naik, 2009), the results for each style are similar to the full sample results.

1.7. The predictive power of a hedge fund's past BAB exposure

In this section, we consider whether a hedge fund's past BAB exposure itself can predict future performance. Titman and Tiu (2011) show that relatively high skill hedge fund managers rely less on factor exposures, and Sun, Wang, and Zheng (2012) find high skill managers also choose to pursue comparatively unique strategies. Therefore, the choice a hedge fund manager makes about whether or not to rely on the BAB factor to generate returns could be a signal of skill. If BAB exposure is a signal of skill, we would expect hedge funds that have had large, positive exposures to the BAB factor in the past to subsequently underperform funds that have had very limited past exposures (assuming the model comparing their performance does not treat BAB exposure as alpha).

We form two portfolios to test the predictive power of BAB exposure. The 'Top BAB' portfolio is an equal weight portfolio of all hedge funds in the highest decile of past BAB exposure, and the 'Btm BAB' portfolio is an equal weight portfolio of all hedge funds in the lowest decile. BAB exposure for each fund is re-estimated at the beginning of each month using OLS regression, the past 24 monthly returns, and the FH8+BAB model. We employ OLS regression rather than LASSO regression to force a measure of BAB exposure for each fund. Instead of using the point estimate of BAB exposure to form the portfolios, we use the *t*-statistic associated with that estimate. This choice lessens the importance of hedge funds with large, but statistically unreliable, point estimates of BAB exposure.

Before testing the performance of the portfolios, we first evaluate whether the hedge funds in them differ across other dimensions. Table 1.5 reports the mean characteristics for the funds in the 'Top BAB' portfolio and the 'Btm BAB' portfolio. As shown, the two groups of funds are not statistically different in many respects, such as assets and hurdle rate; in addition, the majority of the differences that are statistically significant are of limited economic importance. For example, the hedge funds in the 'Top BAB' portfolio have a statistically lower average incentive fee of 19.23%, but the average incentive fee of those in the 'Btm BAB' portfolio is just 19.96%. The similarity of the hedge funds in the two portfolios across these dimensions should provide at least some comfort that any differences identified in portfolio performance are related to BAB exposure, not other fund characteristics.

Table 1.6 shows the performance of the 'Top BAB' portfolio and the 'Btm BAB' portfolio using LASSO regressions and both the FH8 model and the FH8+BAB model.[33] Exposure to the BAB factor is persistent: the BAB factor is not selected by the LASSO regression for the 'Btm BAB' portfolio, but it is selected by the LASSO regression for the 'Top BAB' portfolio, with an exposure of 0.83 (t-stat = 8.61). The key portfolio with respect to testing our hypothesis is the one which longs the 'Btm BAB' portfolio and shorts the 'Top BAB' portfolio, and the key measure for that portfolio is its FH8+BAB model alpha. That long/short portfolio's alpha using that model is 0.39% per month (t-stat = 2.29).[34] Repeating the analysis using OLS regressions in Table 1.7, we find that, for the same portfolio, the same alpha is 0.49% per month (t-stat = 2.77).

[33] As explained before, statistically comparing the alphas for the 'Btm BAB' portfolio across the two models or the long/short portfolio across the two models is either unnecessary or not feasible when using LASSO regressions. We therefore made the choice to keep the table internally consistent and also not tabulate the statistical comparison for the "Top BAB" portfolio.

[34] The LASSO regressions select different sets of factors for the 'Top BAB', 'Btm BAB', and long/short portfolios, so the alpha of the long/short portfolio is not equal to the difference between the alphas of the 'Top BAB' and 'Btm BAB' portfolios.

Hence, when the model does not treat BAB exposure as alpha, hedge funds that have had large, positive exposures to the BAB factor in the past do—to an economically and statistically significant degrees—subsequently underperform funds that have had very limited past exposures. That result supports our hypothesis that a hedge fund manager choosing to use a substantial amount of BAB exposure is a negative signal about their skill (and vice versa).

Alternatively, that result could, perhaps, be dismissed as mechanical: the BAB factor has a positive average return and hedge funds' exposures to the BAB factor are persistent, so we should expect the 'Top BAB' portfolio to underperform the 'Btm BAB' portfolio when the performance model does not treat BAB exposure as alpha. We disagree with this reasoning. While we should expect and do find that the alpha of the 'Top BAB' portfolio decreases more than the alpha of the 'Btm BAB' portfolio when the performance model does not treat BAB exposure as alpha, that expectation does not immediately lead to the expectation that the 'Top BAB' portfolio should, in that case, underperform the 'Btm BAB' portfolio. If we set aside our earlier conjectures regarding the hedge fund industry, there is no reason that the funds in the 'Top BAB' portfolio could not (i) outperform when treating BAB exposure as alpha and (ii) outperform to a lesser extent when not treating BAB exposure as alpha. Put another way, there is no mechanical requirement that precludes hedge funds from having both large exposures to exotic factors and highly skilled management, such that outperformance would be found even if the performance model does not treat exotic exposures as alpha.

1.8. Hedge funds' past BAB exposures and BAB factor timing

Past research has found that many hedge funds can time both the overall market (e.g., Chen and Liang, 2007) and specific factors (e.g., Cao, Chen, Liang, and Lo, 2013). Given those results, we consider in this section whether hedge funds can successfully time the BAB factor. In particular, we focus on the possibility that the hedge funds in the ‘Top BAB’ portfolio are not simply maintaining a constant large exposure to the BAB factor, but rather successfully timing the BAB factor. If such timing is occurring, then our previous conclusions about the relatively low skill of hedge fund managers with large BAB exposures could be erroneous.

Our analysis with respect to this possibility is broken into two parts. We first consider whether the hedge funds in the ‘Top BAB’ or ‘Btm BAB’ portfolio successfully time the BAB factor and re-evaluate their performance under the assumption that any successful timing is true alpha. Then, we investigate the extent to which those hedge funds’ BAB factor timing is achieved using public information. That analysis examines the relation between the BAB exposures of the ‘Top BAB’ and ‘Btm BAB’ portfolios and the lagged TED spread—a common measure of the tightness of leverage constraints that Frazzini and Pedersen (2014) show predicts subsequent BAB factor returns. Following Ferson and Schadt (1996), this later test does not treat returns from timing as alpha if the timing is based on public information, since investors could themselves replicate such a strategy without the aid or expense of a manager.

1.8.1. The performance of the BAB portfolios after accounting for factor timing

So far, neither the FH8 model nor the FH8+BAB model has allowed for factor timing. Our conclusions about the relatively low skill of the managers of the hedge funds in the ‘Top BAB’ portfolio could be an artifact of that constraint. Here, we test that possibility by re-evaluating the ‘Top BAB’ and ‘Btm BAB’ portfolios using a model that both allows for factor timing and treats successful factor timing as alpha.[35]

To capture factor timing, we add Henriksson and Merton (1981) style measures to the FH8+BAB model:

[35] If we evaluate the equal weight portfolio of all hedge funds, the ‘Top α’ portfolio, or the ‘Btm α’ portfolio using the model discussed in this section, we find no evidence of successful timing of the BAB factor.

$$r - r_f = \alpha' + \sum_{i=1}^{8} \beta_i FH_i + \beta_9 BAB + \sum_{i=1}^{8} \gamma_i Max(0, FH_i) + \gamma_9 Max(0, BAB) + \varepsilon \qquad (1.3)$$

In this model, each β_i captures the base factor exposure and each γ_i captures the factor timing. A positive γ_9 suggests successful timing of the BAB factor because it indicates that a fund's BAB exposure is larger in periods when the BAB factor return is positive.[36] α' is the average return from all non-timing related active management (i.e., the alpha unrelated to timing).[37] We capture the average return from timing the Fung and Hsieh (FH) factors as:

$$r_{FH\ Timing} = \sum_{i=1}^{8} \gamma_i \frac{\sum_{t=1}^{N} Max(0, FH_{i,t})}{N} \qquad (1.4)$$

We likewise capture the average return from timing the BAB factor as:

$$r_{BAB\ Timing} = \gamma_9 \frac{\sum_{t=1}^{N} Max(0, BAB_t)}{N} \qquad (1.5)$$

To determine the average total return from active management, we calculate:

$$r_{active} = \alpha' + r_{FH\ Timing} + r_{BAB\ Timing} \qquad (1.6)$$

Table 1.8 shows the performance of the 'Top BAB' and 'Btm BAB' portfolios using LASSO regressions and this timing model. Hedge funds in the 'Top BAB' portfolio appear to generate significant returns from timing the BAB factor. The exposure of the 'Top BAB' portfolio to the BAB timing factor is 0.69 (*t*-stat = 2.51), which generates an average return of 0.53% per month. The 'Top BAB' portfolio, however, still has an average total return from active management of only −0.10% per month (*t*-stat = −0.73). In comparison, despite generating no

[36] Our estimates of factor timing in this analysis could be biased downward for all tested portfolios. Goetzmann, Ingersoll, and Ivkovic (2000) show that Henriksson-Merton style measures are biased downward if monthly data is used but the portfolio manager is making timing decisions on a more frequent basis, which Patton and Ramadoria (2013) show is the case for hedge funds.

[37] All costs (e.g., management fees and trading expenses) are charged to α'. As a result, our estimate of α' should be biased downward, while our estimates of the average returns from timing related active management should be biased upward. Nevertheless, the average total return from active management should have a net bias of zero.

returns from timing the BAB factor, the 'Btm BAB' portfolio has an average total return from active management of 0.32% per month (*t*-stat = 2.37). The difference in average total return from active management between the portfolios of 0.41% per month (*t*-stat = 2.38) is consistent with the difference in alpha of 0.39% per month (*t*-stat = 2.29) shown in Table 1.6. We repeat this analysis using OLS regressions in Table 1.9 and find similar, albeit statistically weaker, results.

Considering the results in Tables 1.8 and 1.9 together, although hedge funds in the 'Top BAB' portfolio do successfully time the BAB factor, treating all of that successful timing as alpha does not change our conclusion that a hedge fund manager choosing to use a substantial amount of BAB exposure is a negative signal about their skill.

1.8.2. Timing the BAB factor using publicly available information

We next consider the extent to which the BAB factor timing of the 'Top BAB' and 'Btm BAB' portfolios is driven by public information and how accounting for public-information-based timing impacts our estimates of portfolio performance. To enable this analysis, we build on Frazzini and Pedersen (2014), who show a relation between the return on the BAB factor and funding constraint tightness. In particular, they show a negative relation between the BAB factor return and the lagged level of the TED spread, which is the difference between the three-month LIBOR and the three-month U.S. Treasuries bill rate.[38,39] Following the logic of Ferson and Schadt (1996), we use that relation to generate a model that (i) identifies timing that can be attributed to the public information about future BAB factor returns embedded in the lagged TED spread and (ii) does not treat BAB factor timing attributable to that information as alpha.

[38] Frazzini and Pedersen (2014) also consider the relation between the contemporaneous TED spread and the return on the BAB factor, but that relation is not investable, so we do not consider it.

[39] In our sample period, we also find a negative relation between the lagged level of the TED spread and the BAB factor return: the average BAB factor return is 0.20% per month when the lagged TED spread is equal to or above its median level, compared to 0.72% per month when its below median.

We begin our analysis by dividing our time period of study into two states—'High TED' and 'Low TED'—depending on whether the one-month lagged level of the TED spread is above or below its sample median.[40] We then evaluate the BAB exposures of the 'Top BAB' and 'Btm BAB' portfolios in each of the states using the FH8+BAB model and both LASSO regressions and OLS regressions. For a given portfolio, if the BAB exposure is greater in the 'Low TED' state compared to the 'High TED' state, that suggests that the hedge funds in the portfolio are timing the BAB factor using public information. If public-information-based timing is occurring, our split design will not treat the elevated BAB exposure in the 'Low TED' state as alpha.

Table 1.10 reports the results for the 'Top BAB' and 'Btm BAB' portfolios. The pattern of the BAB exposures for the 'Top BAB' portfolio is consistent with the hedge funds in that portfolio timing the BAB factor based on public information. Using the LASSO regressions, the 'Top BAB' portfolio has a BAB exposure of 1.41 (*t*-stat = 8.66) in the 'Low TED' state, but an exposure of only 0.57 (*t*-stat = 4.64) in the 'High TED' state. The statistical significance of the difference between those two BAB exposures cannot be assessed because the factors selected by the LASSO regressions change between time periods, so we turn to the OLS regressions for further analysis. Using the OLS regressions, the difference in BAB exposures between the two states is both economically and statistically significant at 0.72 (*t*-stat = 3.22). Conversely, for the 'Btm BAB' portfolio, there is some evidence of timing based on public information—the LASSO regressions select the BAB factor as relevant in the 'Low TED' state, but not the 'High TED' state—but using the OLS regressions, the difference in BAB exposures between the two states is not near statistically significant at any conventional level (*t*-stat = 0.93).

[40]Data on the TED spread is available from the Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/TEDRATE

Comparing the ‘Top BAB’ and ‘Btm BAB’ portfolios in each state, the ‘Btm BAB’ portfolio performs better regardless of the level of the lagged TED spread. When public-information-based BAB factor timing is not treated as alpha, untabulated OLS regressions show that the ‘Btm BAB’ portfolio outperforms the ‘Top BAB’ portfolio by 0.55% per month (*t*-stat = 1.90) in the ‘High TED’ state and 0.64% per month (t-stat = 2.37) in the ‘Low TED’ state. Although, we should consider that result cautiously for three reasons. First, we did not control for any other public information that could help predict the return on the BAB factor or other factors in the model.[41] Second, we did not adapt the model to try to capture as alpha any timing related to nonpublic information. Third, the point in time when it is reasonable to consider the relation between the lagged TED spread and the BAB factor return to be public is unclear. Before the relation was public, the argument that taking advantage of the relation should not be considered alpha is weaker.

Nonetheless, those caveats do not meaningfully affect our conclusion that using substantial amounts of BAB exposure is a signal of relatively low manager skill. The model in this subsection is, perhaps, biased against the ‘Top BAB’ portfolio; however, the model in the prior subsection (8.1) is biased towards the ‘Top BAB’ portfolio and still finds the ‘Btm BAB’ portfolio performs better. The difference in performance between the two portfolios holds even if we switch to performance measures that embrace the idea that investors do not draw a strong distinction between returns from active management and returns from factor exposures (Agarwal, Green, and Ren, 2018). In fact, the difference in performance between the ‘Top BAB’ portfolio and the ‘Btm BAB’ portfolio holds even if all returns are treated equally. The ‘Top BAB’ portfolio has an

[41] Boguth and Simutin (2018) show that BAB factor returns can be predicted by variation in leverage constraint tightness indicated by changes in the average beta of the stocks held by mutual funds, which the findings of Shive and Yun (2013) suggest could be accounted for by hedge funds. Furthermore, Baker and Wurgler (2006) and Hong and Sraer (2016) suggest that measures of investor sentiment and investor disagreement could predict BAB factor returns.

average excess return of 4.3% per year with a Sharpe ratio of 0.45, while the ‘Btm BAB’ portfolio has an average excess return of 6.3% per year with a Sharpe ratio of 0.64. The differences in those measures once again support the idea that relatively low skill hedge fund managers are the most reliant on BAB exposure.

1.9. Conclusion

Past research has focused on the relation between the low beta anomaly and mutual funds. The conditions of the mutual fund industry cause most mutual funds to not capitalize on the anomaly. However, operating under a different set of conditions, hedge funds are well positioned to capitalize. In this paper, we provide the first extensive empirical evidence showing that hedge funds do tend to “bet against beta.” We further show that the performance of the average hedge fund is significantly overstated by not accounting for that tendency and that the tendency is strongest among hedge fund managers with relatively low skill.

Table 1.1: The performance of an equal weight portfolio of all hedge funds

This table shows the net-of-fee performance of an equal weight portfolio of all hedge funds from January 1996 through March 2012. The portfolio is evaluated using both the FH8 model and the FH8+BAB model. Columns (1) and (2) report results from the LASSO regressions, and columns (4) and (5) report results from the OLS regressions. Column (3) tests the differences between columns (1) and (2), while column (6) tests the differences between columns (4) and (5). A "." in the table means that the corresponding factor is not selected by the LASSO regression. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	FH8 LASSO (1)	FH8+BAB LASSO (2)	(2) − (1) (3)	FH8 OLS (4)	FH8+BAB OLS (5)	(5) − (4) (6)
SP500	0.10***	0.17***	0.07**	0.10***	0.17***	0.07**
	[3.71]	[6.17]	[2.46]	[3.72]	[6.23]	[2.51]
SizeSpread	0.10***	0.11***	0.01	0.10***	0.11***	0.01
	[4.24]	[5.05]	[0.43]	[4.26]	[5.11]	[0.46]
EmergMkt	0.18***	0.16***	-0.02	0.18***	0.16***	-0.02
	[9.40]	[9.00]	[-1.10]	[9.40]	[9.04]	[-1.10]
10Year	0.20***	0.15***	-0.04	0.19***	0.14***	-0.05
	[4.52]	[3.77]	[-1.09]	[4.30]	[3.44]	[-1.18]
CreditSpread	0.17***	0.12***	-0.05	0.17***	0.12***	-0.05
	[3.83]	[2.81]	[-1.28]	[3.88]	[2.92]	[-1.24]
BondTrend	.	.		0.00	0.01	0.00
				[0.69]	[1.28]	[0.53]
FxTrend	0.01**	0.01**	-0.00	0.01**	0.01**	-0.00
	[2.55]	[2.42]	[-0.36]	[2.43]	[2.22]	[-0.43]
ComTrend	0.01**	0.01**	-0.00	0.01**	0.01**	-0.00
	[2.31]	[2.13]	[-0.39]	[2.23]	[1.99]	[-0.44]
BAB		0.40***	0.40***		0.40***	0.40***
		[6.10]	[6.10]		[6.19]	[6.19]
Alpha	0.30***	0.12	-0.19**	0.31***	0.13	-0.19**
	[3.71]	[1.43]	[-2.32]	[3.76]	[1.57]	[-2.27]
Observations	195	195	195	195	195	195
Adjusted R^2	0.75	0.79	0.13	0.75	0.79	0.13

Table 1.2: The performance of an equal weight portfolio of all 'traditional' mutual funds

This table shows the net-of-fee performance of an equal weight portfolio of all 'traditional' mutual funds from January 1996 through March 2012. The portfolio is evaluated using both the FH8 model and the FH8+BAB model. Columns (1) and (2) report results from the LASSO regressions, and columns (3) and (4) report results from the OLS regressions. Column (5) tests the differences between columns (3) and (4). A "." in the table means that the corresponding factor is not selected by the LASSO regression. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	FH8 LASSO (1)	FH8+BAB LASSO (2)	FH8 OLS (3)	FH8+BAB OLS (4)	(4) − (3) (5)
SP500	0.49***	0.49***	0.49***	0.50***	0.01
	[37.35]	[37.35]	[37.16]	[35.25]	[0.91]
SizeSpread	0.20***	0.20***	0.20***	0.20***	0.00
	[17.22]	[17.22]	[17.10]	[17.41]	[0.17]
EmergMkt	0.08***	0.08***	0.08***	0.08***	-0.00
	[9.14]	[9.14]	[9.11]	[8.66]	[-0.40]
10Year	0.18***	0.18***	0.18***	0.17***	-0.01
	[8.80]	[8.80]	[8.42]	[7.92]	[-0.43]
CreditSpread	0.16***	0.16***	0.17***	0.16***	-0.01
	[8.01]	[8.01]	[7.97]	[7.44]	[-0.45]
BondTrend	.	.	0.00	0.00	0.00
			[0.81]	[1.00]	[0.19]
FxTrend	.	.	0.00	0.00	-0.00
			[0.57]	[0.42]	[-0.16]
ComTrend	.	.	-0.00	-0.00	-0.00
			[-0.15]	[-0.31]	[-0.16]
BAB		.		0.08**	0.08**
				[2.25]	[2.25]
Alpha	-0.04	-0.04	-0.04	-0.07*	-0.03
	[-1.09]	[-1.09]	[-0.93]	[-1.70]	[-0.83]
Observations	195	195	195	195	195
Adjusted R^2	0.97	0.97	0.97	0.97	-0.02

Table 1.3: The performance of an equal weight portfolio of all 'alternative' mutual funds

This table shows the net-of-fee performance of an equal weight portfolio of all 'alternative' mutual funds from January 1996 through March 2012. The portfolio is evaluated using both the FH8 model and the FH8+BAB model. Columns (1) and (2) report results from the LASSO regressions, and columns (4) and (5) report results from the OLS regressions. Column (3) tests the differences between columns (1) and (2), while column (6) tests the differences between columns (4) and (5). A "." in the table means that the corresponding factor is not selected by the LASSO regression. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	FH8 LASSO (1)	FH8+BAB LASSO (2)	(2) − (1) (3)	FH8 OLS (4)	FH8+BAB OLS (5)	(5) − (4) (6)
SP500	0.29***	0.34***	0.04***	0.29***	0.34***	0.04***
	[19.50]	[23.11]	[2.97]	[19.69]	[23.10]	[2.87]
SizeSpread	0.12***	0.13***	0.01	0.12***	0.12***	0.01
	[9.15]	[10.82]	[0.49]	[9.08]	[10.69]	[0.52]
EmergMkt	0.08***	0.07***	-0.01	0.08***	0.07***	-0.01
	[8.11]	[7.70]	[-1.32]	[8.09]	[7.68]	[-1.26]
10Year	0.15***	0.12***	-0.03	0.15***	0.12***	-0.03
	[6.45]	[5.85]	[-1.33]	[6.46]	[5.78]	[-1.35]
CreditSpread	0.11***	0.08***	-0.03	0.11***	0.08***	-0.03
	[4.56]	[3.64]	[-1.42]	[4.79]	[3.85]	[-1.42]
BondTrend	.	.		-0.00	-0.00	0.00
				[-0.91]	[-0.42]	[0.61]
FxTrend	.	.		0.01*	0.00*	-0.00
				[1.97]	[1.72]	[-0.50]
ComTrend	.	.		0.00	-0.00	-0.00
				[0.42]	[-0.03]	[-0.51]
BAB		0.25***	0.25***		0.24***	0.24***
		[7.31]	[7.31]		[7.07]	[7.07]
Alpha	-0.08*	-0.20***	-0.12***	-0.09*	-0.19***	-0.11**
	[-1.81]	[-4.66]	[-2.76]	[-1.92]	[-4.60]	[-2.59]
Observations	195	195	195	195	195	195
Adjusted R^2	0.92	0.94	0.20	0.92	0.94	0.17

Table 1.4: The performance of the top and bottom alpha decile portfolios

This table shows the net-of-fee performance of monthly rebalanced portfolios of hedge funds formed based on past fund alpha. At the beginning of each month, we run separate LASSO regressions on the past 24 monthly excess returns of each fund using the FH8+BAB model. We then sort the funds into deciles based on their alphas. The 'Top α' portfolio is an equal weight portfolio of all the funds in the highest alpha decile. The 'Btm α' portfolio is an equal weight portfolio of all the funds in the lowest alpha decile. The period of evaluation for the portfolios is January 1998 through March 2012. The portfolios are evaluated using both the FH8 model and the FH8+BAB model. Columns (1) through (4) report results from using LASSO regressions to evaluate portfolio performance. Columns (5), (6), (8), and (9) report results from using OLS regressions to evaluate portfolio performance. Column (7) tests the differences between columns (5) and (6), and column (10) tests the differences between columns (8) and (9). A "." in the table means that the corresponding factor is not selected by the LASSO regression. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Top α portfolio		Btm α portfolio		Top α portfolio			Btm α portfolio		
	FH8 LASSO	FH8+BAB LASSO	FH8 LASSO	FH8+BAB LASSO	FH8 OLS	FH8+BAB OLS	(6) − (5)	FH8 OLS	FH8+BAB OLS	(9) − (8)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SP500	.	.	0.08	0.21***	0.01	0.06	0.05	0.09	0.22***	0.13**
			[1.28]	[3.48]	[0.20]	[0.79]	[0.61]	[1.48]	[3.56]	[2.10]
SizeSpread	0.37***	0.37***	.	.	0.38***	0.39***	0.01	-0.01	0.02	0.03
	[6.49]	[6.49]			[6.38]	[6.54]	[0.17]	[-0.15]	[0.43]	[0.59]
EmergMkt	0.26***	0.26***	0.28***	0.25***	0.26***	0.25***	-0.01	0.28***	0.24***	-0.04
	[9.57]	[9.57]	[7.20]	[6.76]	[5.79]	[5.43]	[-0.28]	[6.86]	[6.30]	[-0.97]
10Year	.	.	0.28***	.	0.10	0.06	-0.04	0.28***	0.17*	-0.10
			[3.05]		[0.94]	[0.58]	[-0.34]	[2.97]	[1.97]	[-1.16]
CreditSpread	.	.	0.31***	0.15*	0.05	0.01	-0.03	0.36***	0.26***	-0.10
			[3.54]	[1.93]	[0.45]	[0.11]	[-0.33]	[3.90]	[2.98]	[-1.13]
BondTrend	.	.	.	.	0.02	0.02	0.00	0.01	0.01	0.01
					[1.24]	[1.39]	[0.15]	[0.41]	[0.97]	[0.53]
FxTrend	.	.	.	.	0.00	0.00	-0.00	0.01	0.01	-0.00
					[0.25]	[0.16]	[-0.09]	[0.95]	[0.71]	[-0.31]
ComTrend	0.03**	0.03**	.	.	0.03**	0.03*	-0.00	0.02	0.01	-0.01
	[2.47]	[2.47]			[2.02]	[1.89]	[-0.13]	[1.48]	[1.14]	[-0.45]
BAB		.		0.80***		0.25	0.25		0.73***	0.73***
				[5.89]		[1.52]	[1.52]		[5.26]	[5.26]
Alpha	0.74***	0.74***	0.02	-0.27	0.74***	0.63***	-0.11	0.05	-0.26	-0.31*
	[3.84]	[3.84]	[0.11]	[-1.53]	[3.70]	[3.01]	[-0.50]	[0.27]	[-1.46]	[-1.73]
Observations	171	171	171	171	171	171	171	171	171	171
Adjusted R^2	0.53	0.53	0.61	0.66	0.52	0.53	-0.04	0.61	0.67	0.10

Table 1.5: The characteristics of the top and bottom BAB hedge funds

This table shows mean characteristics for hedge funds in the top and bottom deciles of past BAB exposure *t*-statistic. The BAB exposure *t*-statistic is estimated for each fund using an OLS regression, the past 24 months of excess returns, and the FH8+BAB model. The differences in the mean characteristics and the *t*-statistics associated with those differences are also reported. The *t*-statistics are calculated using standard errors clustered on fund and year-month. We trim the top and bottom 1% of all continuous variables. The variables listed after "Redemption Notice" are all indicators, with all the previous variables being continuous. "Directional Traders", "Relative Value", "Security Selection", and "Multiprocess" are broad style categories identified following Agarwal, Daniel, and Naik (2009).

	Top BAB Funds	Btm BAB Funds	Difference	Diff *t*-stat
Assets ($M)	161.17	139.48	21.69	0.85
Age (months)	84.73	80.69	4.03	1.78
Min Invest ($M)	0.53	0.76	-0.23	-1.47
Mgmt Fee (%)	1.47	1.58	-0.11	-2.91
Incentive Fee (%)	19.23	19.96	-0.73	-4.02
Hurdle Rate (%)	5.57	6.12	-0.55	-0.71
Lockup Period (days)	304.81	355.32	-50.51	-1.37
Redemption Notice (days)	41.60	35.03	6.57	1.28
High Water Mark	0.79	0.87	-0.08	-2.54
Offshore (non-US)	0.84	0.64	0.20	4.84
Directional Traders	0.18	0.28	-0.09	-3.14
Relative Value	0.22	0.16	0.06	1.93
Security Selection	0.35	0.37	-0.02	-0.61
Multiprocess	0.20	0.14	0.06	2.29

Table 1.6: The performance of the BAB portfolios using LASSO

This table shows the net-of-fee performance of monthly rebalanced portfolios of hedge funds formed based on past BAB exposure *t*-statistic. At the beginning of each month, we run separate OLS regressions on the past 24 monthly excess returns of each fund using the FH8+BAB model. We then sort the funds into deciles based on their BAB exposure *t*-statistics. The 'Top BAB' portfolio is an equal weight portfolio of all the funds in the highest BAB decile. The 'Btm BAB' portfolio is an equal weight portfolio of all the funds in the lowest BAB decile. The period of evaluation for the portfolios is January 1998 through March 2012. The portfolios are evaluated using LASSO regressions and both the FH8 model and the FH8+BAB model. A "." in the table means that the corresponding factor is not selected by the LASSO regression. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Top BAB portfolio		Btm BAB portfolio		Btm − Top	
	FH8	FH8+BAB	FH8	FH8+BAB	FH8	FH8+BAB
	(1)	(2)	(3)	(4)	(5)	(6)
SP500	.	0.17***	.	.	.	.
		[4.01]				
SizeSpread	.	.	0.33***	0.33***	0.33***	0.33***
			[8.26]	[8.26]	[6.78]	[7.02]
EmergMkt	0.22***	0.17***	0.21***	0.21***	.	.
	[9.45]	[6.53]	[10.87]	[10.87]		
10Year	0.33***	0.22***	.	.	-0.21**	.
	[4.61]	[3.65]			[-2.40]	
CreditSpread	0.30***	0.19***	.	.	-0.23***	-0.17***
	[4.34]	[3.12]			[-3.30]	[-2.86]
BondTrend	.	.	.	.	.	.
FxTrend	.	.	.	.	.	.
ComTrend	.	.	.	.	.	.
BAB		0.83***		.		-0.62***
		[8.61]				[-5.07]
Alpha	0.07	-0.29**	0.29**	0.29**	0.17	0.39**
	[0.50]	[-2.31]	[2.09]	[2.09]	[1.00]	[2.29]
Observations	171	171	171	171	171	171
Adjusted R^2	0.57	0.70	0.61	0.61	0.23	0.31

Table 1.7: The performance of the BAB portfolios using OLS

This table shows the net-of-fee performance of the BAB portfolios defined in Table 1.6 estimated using OLS regressions. The period of evaluation for the portfolios is January 1998 through March 2012. The portfolios are evaluated using both the FH8 model and the FH8+BAB model. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Top BAB portfolio			Btm BAB portfolio			Btm − Top		
	FH8	FH8+BAB	(2) − (1)	FH8	FH8+BAB	(5) − (4)	FH8	FH8+BAB	(8) − (7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
SP500	0.03	0.18***	0.15***	0.10**	0.13**	0.03	0.06	-0.05	-0.12*
	[0.71]	[4.10]	[3.32]	[2.04]	[2.49]	[0.61]	[1.05]	[-0.83]	[-1.85]
SizeSpread	-0.01	0.02	0.03	0.35***	0.35***	0.01	0.36***	0.33***	-0.03
	[-0.26]	[0.62]	[0.93]	[8.38]	[8.54]	[0.17]	[6.85]	[6.71]	[-0.52]
EmergMkt	0.20***	0.16***	-0.04	0.15***	0.14***	-0.01	-0.05	-0.02	0.03
	[6.34]	[5.89]	[-1.54]	[4.82]	[4.48]	[-0.28]	[-1.25]	[-0.44]	[0.86]
10Year	0.35***	0.24***	-0.12*	0.13*	0.10	-0.02	-0.23**	-0.14	0.09
	[4.82]	[3.76]	[-1.84]	[1.73]	[1.36]	[-0.34]	[-2.48]	[-1.54]	[1.03]
CreditSpread	0.31***	0.20***	-0.11*	0.10	0.07	-0.02	-0.21**	-0.13	0.09
	[4.28]	[3.20]	[-1.79]	[1.35]	[0.99]	[-0.33]	[-2.36]	[-1.45]	[1.00]
BondTrend	-0.01	-0.01	0.01	0.00	0.00	0.00	0.01	0.01	-0.01
	[-1.30]	[-0.70]	[0.84]	[0.10]	[0.25]	[0.15]	[1.11]	[0.71]	[-0.47]
FxTrend	0.01	0.01	-0.00	0.01	0.01	-0.00	-0.00	0.00	0.00
	[1.45]	[1.23]	[-0.49]	[1.34]	[1.25]	[-0.09]	[-0.10]	[0.17]	[0.27]
ComTrend	0.01	0.01	-0.01	0.00	0.00	-0.00	-0.01	-0.00	0.01
	[1.08]	[0.57]	[-0.71]	[0.35]	[0.22]	[-0.13]	[-0.59]	[-0.22]	[0.40]
BAB		0.82***	0.82***		0.18	0.18		-0.64***	-0.64***
		[8.32]	[8.32]		[1.52]	[1.52]		[-4.64]	[-4.64]
Alpha	0.05	-0.29**	-0.34***	0.27*	0.20	-0.07	0.22	0.49***	0.27
	[0.38]	[-2.31]	[-2.74]	[1.95]	[1.34]	[-0.50]	[1.24]	[2.77]	[1.53]
Observations	171	171	171	171	171	171	171	171	171
Adjusted R^2	0.57	0.70	0.26	0.62	0.62	-0.04	0.22	0.31	0.07

Table 1.8: The performance of the BAB portfolios using LASSO after accounting for market timing

This table shows the net-of-fee performance of the BAB portfolios defined in Table 1.6 using LASSO regressions after accounting for market timing. The period of evaluation for the portfolios is January 1998 through March 2012. The portfolios are evaluated using the FH8+BAB model with timing factors. The first column for each portfolio reports the base factor exposures, and the second column reports the exposures to the Henriksson and Merton (1981) style timing factors. The bottom rows of the table report the components of each portfolio's active return (FH8 Timing Return, BAB Timing Return, and Non-Timing Alpha) and the Total Active Return, which is the sum of the components. A "." in the table means that the corresponding factor is not selected by the LASSO regression. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Top BAB Portfolio		Btm BAB Portfolio		Btm − Top	
	Factor Exposure	Factor Timing: Max(0, Factor)	Factor Exposure	Factor Timing: Max(0, Factor)	Factor Exposure	Factor Timing: Max(0, Factor)
	(1)	(2)	(3)	(4)	(5)	(6)
SP500	0.19***	.	.	.	.	.
	[4.36]					
SizeSpread	.	.	0.15**	0.33***	0.15*	0.27**
			[2.12]	[2.99]	[1.75]	[2.00]
EmergMkt	0.17***	.	0.21***	.	.	.
	[6.55]		[11.32]			
10Year	0.21***	.	.	.	.	.
	[3.39]					
CreditSpread	0.17***	.	.	.	.	.
	[2.80]					
BondTrend	.	.	.	.	.	.
FxTrend	.	.	.	.	.	.
ComTrend	.	.	.	.	.	.
BAB	0.44**	0.69**	.	.	-0.58***	.
	[2.44]	[2.51]			[-4.75]	
Observations	171		171		171	
Adjusted R^2	0.71		0.63		0.29	
FH8 Timing Return	0.00		0.46		0.38	
BAB Timing Return	0.53		0.00		0.00	
Non-Timing Alpha	-0.64		-0.14		0.04	
Total Active Return	-0.10		0.32**		0.41**	
	[-0.73]		[2.37]		[2.38]	

Table 1.9: The performance of the BAB portfolios using OLS after accounting for market timing

This table shows the net-of-fee performance of the BAB portfolios defined in Table 1.6 using OLS regressions after accounting for market timing. The period of evaluation for the portfolios is January 1998 through March 2012. The portfolios are evaluated using the FH8+BAB model with timing factors. The first column for each portfolio reports the base factor exposures, and the second column reports the exposures to the Henriksson and Merton (1981) style timing factors. The bottom rows of the table report the components of each portfolio's active return (FH8 Timing Return, BAB Timing Return, and Non-Timing Alpha) and the Total Active Return, which is the sum of the components. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Top BAB Portfolio		Btm BAB Portfolio		Btm − Top	
	Factor Exposure (1)	Factor Timing: Max(0, Factor) (2)	Factor Exposure (3)	Factor Timing: Max(0, Factor) (4)	Factor Exposure (5)	Factor Timing: Max(0, Factor) (6)
SP500	0.11	0.12	0.10	0.04	-0.01	-0.08
	[1.59]	[1.16]	[1.20]	[0.31]	[-0.08]	[-0.53]
SizeSpread	-0.05	0.12	0.16**	0.35***	0.20**	0.23*
	[-0.72]	[1.21]	[2.04]	[2.99]	[2.21]	[1.67]
EmergMkt	0.10**	0.13*	0.15***	-0.00	0.06	-0.13
	[2.10]	[1.88]	[2.79]	[-0.03]	[0.89]	[-1.32]
10Year	0.36***	-0.23	0.24	-0.27	-0.12	-0.05
	[3.03]	[-1.13]	[1.65]	[-1.12]	[-0.71]	[-0.16]
CreditSpread	0.41***	-0.36**	0.01	0.16	-0.40***	0.52**
	[3.95]	[-2.39]	[0.12]	[0.89]	[-2.62]	[2.39]
BondTrend	-0.01	0.02	0.01	-0.01	0.03	-0.04
	[-0.79]	[0.92]	[0.54]	[-0.41]	[1.00]	[-0.98]
FxTrend	0.04**	-0.05**	0.02	-0.01	-0.02	0.04
	[2.13]	[-2.10]	[0.76]	[-0.31]	[-0.83]	[1.19]
ComTrend	0.01	-0.01	0.02	-0.03	0.02	-0.02
	[0.35]	[-0.43]	[0.91]	[-0.91]	[0.52]	[-0.47]
BAB	0.37*	0.74**	0.15	0.09	-0.22	-0.65
	[1.68]	[2.38]	[0.55]	[0.25]	[-0.69]	[-1.43]
Observations	171		171		171	
Adjusted R^2	0.73		0.63		0.33	
FH8 Timing Return	0.01		0.16		0.15	
BAB Timing Return	0.58		0.07		-0.50	
Non-Timing Alpha	-0.63		0.01		0.64	
Total Active Return	-0.04		0.24		0.28	
	[-0.27]		[1.27]		[1.25]	

Table 1.10: The TED spread and the BAB exposures of the BAB portfolios

This table shows the net-of-fee performance of the BAB portfolios defined in Table 1.6 conditional on the lagged level of the TED spread. The TED spread is the difference between the three-month LIBOR and the three-month U.S. Treasury bill rate. The period of evaluation for the portfolios is January 1998 through March 2012. Columns (1), (3), (5), and (8), which are labeled 'High TED', report the results for months in which the lagged by one month TED spread is higher than or equal to the sample median. Columns (2), (4), (6), and (9), which are labeled 'Low TED', report the results for the other months. The portfolios are evaluated using the FH8+BAB model. Columns (1) through (4) report results from using LASSO regressions to evaluate portfolio performance. Columns (5), (6), (8), and (9) report results from using OLS regressions to evaluate portfolio performance. Column (7) tests the differences between columns (5) and (6), and column (10) tests the differences between columns (8) and (9). A "." in the table means that the corresponding factor is not selected by the LASSO regression. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Top BAB portfolio		Btm BAB portfolio		Top BAB portfolio			Btm BAB portfolio		
	High TED	Low TED	High TED	Low TED	High TED	Low TED	(5) − (6)	High TED	Low TED	(8) − (9)
	LASSO	LASSO	LASSO	LASSO	OLS	OLS		OLS	OLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SP500	0.14***	0.28***	.	0.16**	0.15***	0.30***	-0.15	0.15*	0.15**	0.00
	[2.75]	[3.96]		[2.57]	[2.73]	[3.65]	[-1.47]	[1.94]	[2.43]	[0.01]
SizeSpread	.	.	0.41***	0.20***	0.02	0.07	-0.06	0.46***	0.20***	0.26***
			[7.10]	[4.05]	[0.35]	[1.10]	[-0.71]	[7.35]	[4.01]	[2.81]
EmergMkt	0.18***	0.15***	0.21***	0.13***	0.17***	0.12**	0.05	0.15***	0.13***	0.02
	[5.74]	[3.09]	[7.83]	[3.17]	[5.16]	[2.18]	[0.83]	[3.27]	[3.15]	[0.33]
10Year	0.21**	.	.	0.20***	0.22**	0.22**	0.00	-0.08	0.21***	-0.29*
	[2.43]			[3.13]	[2.40]	[2.54]	[0.01]	[-0.64]	[3.10]	[-1.95]
CreditSpread	0.24***	.	.	0.20**	0.25***	0.12	0.12	-0.01	0.20**	-0.21
	[3.21]			[2.54]	[3.12]	[1.14]	[0.93]	[-0.11]	[2.46]	[-1.35]
BondTrend	.	.	.	0.01	-0.00	-0.00	-0.00	0.02	-0.00	0.02
				[1.46]	[-0.25]	[-0.02]	[-0.16]	[1.22]	[-0.22]	[1.14]
FxTrend	.	.	.	.	0.01	0.00	0.01	0.01	0.01	-0.00
					[0.77]	[0.10]	[0.47]	[0.65]	[1.35]	[-0.06]
ComTrend	.	.	.	.	-0.00	0.02	-0.02	-0.00	0.00	-0.01
					[-0.26]	[1.24]	[-1.06]	[-0.16]	[0.32]	[-0.29]
BAB	0.57***	1.41***	.	0.37***	0.59***	1.30***	-0.72***	0.12	0.36***	-0.24
	[4.64]	[8.66]		[2.90]	[4.58]	[7.32]	[-3.22]	[0.65]	[2.65]	[-0.93]
Alpha	-0.18	-0.66***	0.34	-0.06	-0.19	-0.70***	0.51*	0.36	-0.06	0.42
	[-1.11]	[-3.20]	[1.48]	[-0.40]	[-1.12]	[-3.26]	[1.86]	[1.54]	[-0.37]	[1.32]
Observations	87	84	87	84	87	84		87	84	
Adjusted R^2	0.76	0.63	0.63	0.72	0.74	0.64		0.62	0.72	

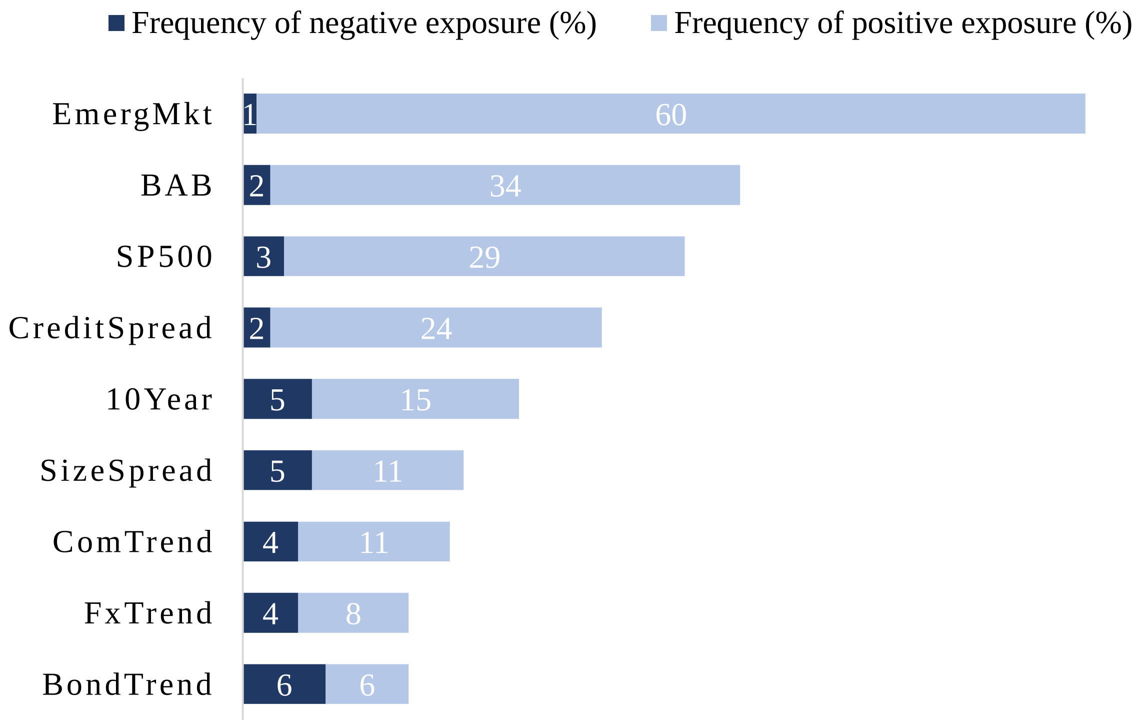

Figure 1.1: The frequency of the BAB and FH8 factors being selected by LASSO regressions
This figure shows how often each factor in Eq. (1.2) is selected across 6,163 fund-level LASSO (least absolute shrinkage and selection operator) regressions. We require each hedge fund to have at least 24 monthly observations and run LASSO regressions on the monthly excess net returns of each hedge fund using the FH8 (Fung and Hsieh eight) factors and the BAB (betting-against-beta) factor. The full sample period is January 1996 through March 2012, but the period for each individual fund is a subperiod of the full sample period.

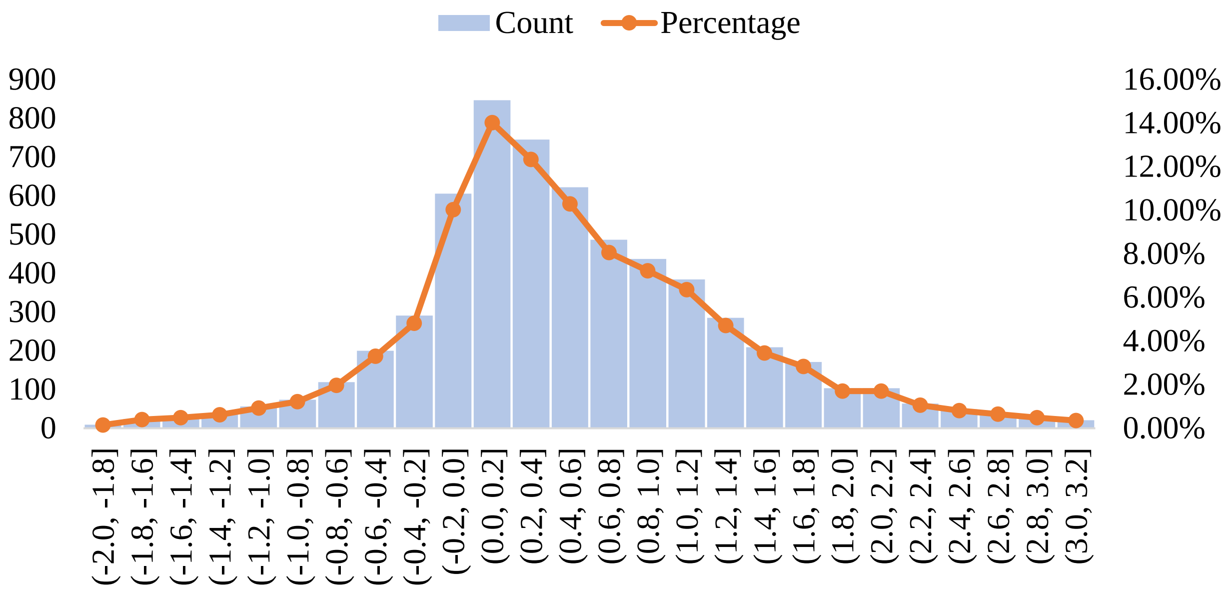

Figure 1.2: The distribution of hedge funds' BAB exposures

This figure shows the distribution of BAB (betting-against-beta) exposures resulting from 6,163 fund-level OLS regressions of monthly excess net fund returns on the FH8 (Fung and Hsieh eight) and BAB factors. We require each hedge fund to have at least 24 observations and trim the top 1% and bottom 1% of BAB exposures. The full sample period is January 1996 through March 2012, but the period for each individual fund is a subperiod of the full sample period.

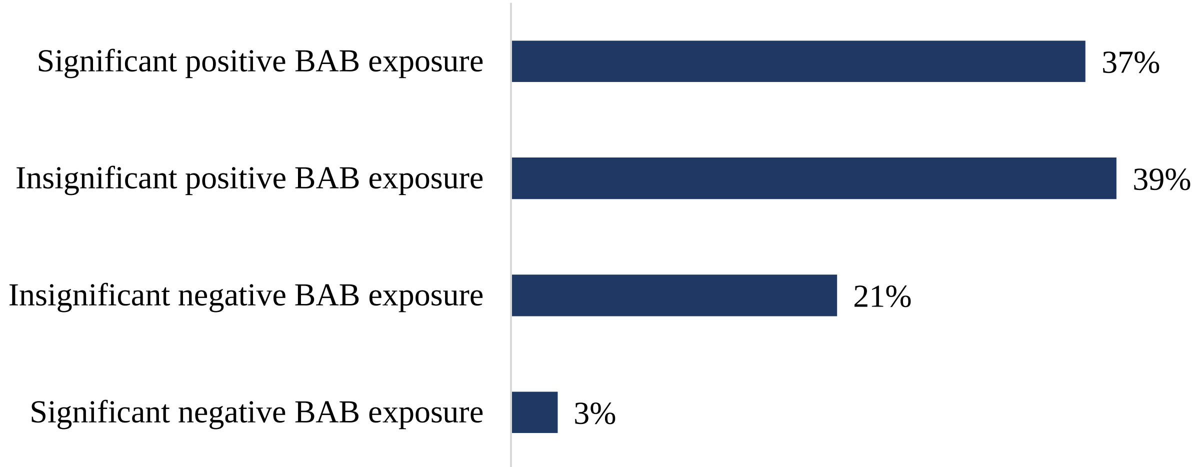

Figure 1.3: The statistical significance of hedge funds' BAB exposures
This figure shows the distribution of BAB (betting-against-beta) exposures by sign and statistical significance at the 10% level. The percentages reported are derived from 6,163 fund-level OLS regressions of monthly excess net fund returns on the FH8 (Fung and Hsieh eight) and BAB factors. We require each hedge fund to have at least 24 observations. The full sample period is January 1996 through March 2012, but the period for each individual fund is a subperiod of the full sample period.

Chapter 2. Hedge Fund Performance Prediction with Machine Learning

Alexey Malakhov, Timothy B. Riley, Qing Yan

2.1. Introduction

The popular press and academic research both indicate that, over the past decade, the performance of the hedge fund industry has declined.[42] Despite that trend, hedge funds are still an important component of the financial market with more than $5 trillion in assets under management.[43] Given their prominence, there are significant gains to being able to identify the hedge funds that can be expected to outperform. Many hedge fund performance predictors have been proposed in literature; however, substantial uncertainty remains. Do the predictors still have power? Which predictors are most important? Can the predictors be effectively used in combination? In the absence of a framework to analyze these questions, the literature has been focused on either individual novel predictors or ad hoc combinations of predictors.[44]

In this study, we use multiple machine learning models to dynamically identify and combine 22 previously documented predictors of hedge fund performance. We find that these models are able to identify outperforming hedge funds in both the full sample and key subperiods.

We train two linear models (OLS and LASSO) and three tree-based nonlinear models (bagged regression trees, random forest, and boosted regression trees) to learn the relation between those 22 predictors and two-year-ahead alpha. We then apply the fitted models out-of-sample and evaluate their predictive power. The models are re-fitted and predictions are made from 2002 to

[42] See, for example, "A losing bet: Hedge funds haven't delivered on their promise," The Economist, May 7th, 2016 and Bollen, Joenväärä, and Kauppila (2020).
[43] Using the most comprehensive data available, Barth, Joenväärä, Kauppila, and Wermers (2019) show that the net assets under management for hedge fund industry were, in 2016, more than $5.2 trillion.
[44] See, for example, Fung and Hsieh (1997, 2001, 2004); Kosowski, Naik, and Teo (2007); Bali, Gokcan, and Liang (2007); Agarwal, Daniel, and Naik (2009); Patton (2009); Sadka (2010); Jagannathan, Malakhov, and Novikov (2010); Titman and Tiu (2011); Sun, Wang, and Zheng (2011); Cao, Chen, Liang, and Lo (2013); Bali, Brown, and Caglayan (2011, 2012, 2014, 2019); and Duanmu, Malakhov, and McCumber (2018).

2017. Over that period, a one-dollar investment in an equal weight, annually rebalancing portfolio of the hedge funds with the highest predicted alphas from the linear models returns $7.03. Using the non-linear models, the return is $5.21. In comparison, an equal weight portfolio of all hedge funds returns just $2.50, and an investment in the S&P 500 returns just $3.22.

The portfolio of hedge funds with the highest predicted alphas from each of the five machine learning models delivers economically large and statistically significant out-of-sample alpha. The alphas range from 5.16% per year (*t*-stat = 2.24) to 7.80% per year (*t*-stat = 3.52). During the 2008 to 2017 subperiod—during which Bollen, Joenväärä, and Kauppila (2020) find allocations to hedge funds do not improve risk-adjusted performance—the portfolios formed based on the machine learning models still deliver superior performance. For example, the portfolio generated by the LASSO model during that period has an alpha of 8.88% per year (*t*-stat = 2.91), an annual return of 11.64%, and a Sharpe ratio of 0.80. The S&P 500 over the same time period had a 9% annual return and a Sharpe ratio of 0.59 (and, by construction, an alpha of zero).

Our models are designed to allow the importance of each predictor to vary over time. However, we find that average return and alpha; Bali, Brown, and Caglayan (2012) systematic risk; Duanmu, Malakhov, and McCumber (2018) beta activity; and Bali, Brown, and Caglayan (2019) maximum return are consistently important. The alpha *t*-statistic, the Sharpe ratio, and Titman and Tiu (2011) R^2 are, on average, less important.

To compare with the traditional approach, we also form portfolios based on the individual predictors. None of the resulting 22 portfolios has a greater alpha than the LASSO-based machine learning portfolio. Further, the average alpha of the portfolios based on the individual predictors is 3.72% per year, which is 28% lower than the average across the machine learning portfolios (5.16%). The individual portfolios tend to perform well when the predictor is aligned with what is being predicted (e.g., using past Sharpe ratio to predict future Sharpe ratio), whereas the machine learning portfolios perform well across all common metrics.

The machine learning models provide information about future hedge fund performance that is not captured by the individual predictors. We isolate a given machine learning model's unique information by excluding from that model's portfolio any hedge funds that both the machine learning model and an individual predictor identify as expected future outperformers. After this exclusion, the machine learning portfolios all retain their superior out-of-sample performance. Conversely, if the same method is applied to isolate the individual predictors' unique information, the individual predictor portfolios tend to perform significantly worse. Hence, only the machine learning models generate unique information of value to investors.

There is a substantial literature on hedge fund performance prediction, but the prior work tends to focus on individual predictors. The study most closely related to ours is Bollen, Joenväärä, and Kauppila (2020), which also considers a set of predictors. They consider 26 predictors and use hedge funds' ranks with respect to those predictors to perform sorts—implicitly assuming time invariant, equal predictive power for each measure. Our machine learning models, on the other hand, allow predictive power to vary in both the time-series and cross-section. This allowance is important as investor learning (McLean and Pontiff, 2016) and changing market conditions (Kacperczyk, van Nieuwerburgh and Veldkamp, 2014) should, in expectation, cause a given predictor's power to be nonconstant. The relaxation of assumptions about predictive power that machine learning permits could explain the markedly different results between our study and Bollen, Joenväärä, and Kauppila (2020) with respect to hedge fund performance from 2008 onward.

The contention of Harvey, Liu, and Zhu (2016, pg. 5) that "most claimed research findings in financial economics are likely false" has generated a surge of re-evaluations in the financial literature, with the asset management space being no exception (e.g., Riley, 2019, and Jones and Mo, 2019). We do not directly replicate any prior results—either in- or out-of-sample—and our analysis is not immune from bias, but our analysis does provide a framework for the parsimonious simultaneous re-evaluation of a set of predictors. With our framework, it is possible to efficiently separate the predictors that do have unique predictive power from those predictors that do not, while also capturing information about when a given predictor has been useful.

Despite its wide application in general, machine learning is sparsely applied in financial research. Khandani, Kim and Lo (2010) use machine learning models to forecast credit card delinquencies and defaults; Rapach, Strauss, and Zhou (2013) and Gu, Kelly, and Xiu (2020) use machine learning models to predict stock returns; and Culkin and Das (2017) use machine learning models to predict option prices. Our study, being the first to use machine learning models to predict hedge fund performance, contributes a novel and economically important machine learning application to this growing literature.

2.2. Data and methods

2.2.1. Hedge fund data

We collect the hedge fund data from Bloomberg database over the period January 1994 – December 2018. The data includes monthly net returns and common fund characteristics, such as assets under management, fees, and style.[45] We start with 20,072 unique funds, including both live and defunct funds. Like other databases, hedge funds report data to Bloomberg on a voluntary basis, although Bloomberg requires that all reporting funds provide performance since inception. To reduce backfill bias, we drop the first 24 months of each fund’s returns, which remains 13,883

[45] Hedge fund returns are calculated from fund net asset values (NAVs) converted into U.S. dollars.

funds.[46] For training dataset, we require four years return history to calculate all the predictors and two-year-ahead returns to calculate FH8 alpha. For out-of-sample testing dataset, we require four years return history to have all the predictors available. This leads to a final sample of 3,820 funds in the training dataset and 5,137 funds in the testing dataset.

Panel A of Table 2.1 provides summary statistics of all 20,072 hedge funds on fund returns, assets, fees, and fund longevity. Since medians are less affected by outliers and thus more representative, we focus on the medians. The representative fund has a 0.39% average excess monthly return, $45.4 million under management, a 1.5% management fee, and a 20% performance fee.

Panel B of Table 2.1 reports the number of unique funds and the average return of an equal weight portfolio of all hedge funds by year. The highest average monthly return of 2.08% (24.96% per year) in 1999 is consistent with the results in Joenväärä, Kauppila, Kosowski, and Tolonen (2019). The lowest average monthly return of –1.61% (–19.32% per year), and highest monthly volatility of 3.59% (12.44% per year) both occur in the 2008 financial crisis.

2.2.2. Fund performance measure and predictors

In this study, we train five machine learning models to predict two-year-ahead Fung and Hsieh (2004) eight-factor model alpha (FH8 thereafter) with 22 predictors.[47] Although there are studies suggest certain fund characteristics are associated with fund performance, those characteristics seldom change over time.[48] Given the dynamic feature of hedge funds, we use return based variables instead of fund characteristics as predictors. We consider three simple return

[46] The date upon which a hedge fund first reports data to Bloomberg is unavailable. Therefore, we do not know the exact number of backfilled returns for any given fund. The 24-month correction is consistent with both Jagannathan, Malakhov, and Novikov (2010) and Titman and Tiu (2011).
[47] We thank David Hsieh for making the trend-following factor data available at his website. https://faculty.fuqua.duke.edu/~dah7/HFRFData.htm
[48] For example, Aragon (2007) and Agarwal, Daniel, and Naik (2009) both find funds with lockup restrictions deliver better performance.

based measures: average return, Sharpe ratio, and MAX of Bali, Brown, and Caglayan (2019). We also consider five regression based measures from FH8 model: alpha, *t*-statistic of alpha, R^2 of Titman and Tiu (2011), systematic risk of Bali, Brown, and Caglayan (2013) and beta activity of Duanmu, Malakhov, and McCumber (2018). Prior studies usually focus on one past horizon to calculate predictors. However, since machine learning models can sort out the most relevant predictors and therefore allow us to flexibly consider multiple horizons for the same predictive variable. We include up to four different horizons for a give variable.[49] This leads to a total number of 22 predictors:

(7) Fund average return (Ret_1year), Sharpe ratio (Sharpe_1year), and MAX using fund past one year's fund returns;

(8) Fund average return (Ret_2year), Sharpe ratio (Sharpe_2year), FH8 alpha (Alpha_2year), *t*-statistic of FH8 alpha (Talpha_2year), R^2 (R2_2year), and systematic risk (Sys.Risk_2year) using fund past two years' fund returns;

(9) Fund average return (Ret_3year), Sharpe ratio (Sharpe_3year), FH8 alpha (Alpha_3year), *t*-statistic of FH8 alpha (Talpha_3year), R^2 (R2_3year), and systematic risk (Sys.Risk_3year) using fund past three years' fund returns;

(10) Fund average return (Ret_4year), Sharpe ratio (Sharpe_4year), FH8 alpha (Alpha_4year), *t*-statistic of FH8 alpha (Talpha_4year), R^2 (R2_4year), systematic risk (Sys.Risk_4year) and beta activity (Beta Activity) using fund past four years' fund returns.

[49] To reduce measurement error, we do not include regression-based variables using one year's returns. Also, to avoid extremely high correlation, we only include MAX from the past one year's returns. Beta Activity measure uses past four years' returns by design.

2.2.3. Training and prediction

The general machine learning framework in this study is to train different models to learn the relation between the 22 predictors and two-year-ahead FH8 alpha, and then apply the fitted model to a testing dataset for predictions.

Specifically, to predict two-year-ahead FH8 alpha at the beginning of year t, we fit different models with the training dataset which is a cross-sectional dataset with the FH8 alpha calculated from year t-1 and year t-2 and 22 predictors for each fund calculated using four years' data from year t-6 to year t-3. For example, since our full sample starts from January 1996 and ends in December 2018, the first training set uses the data from 1996 to 2001. We fit different models with the 22 predictors calculated using 1996 to 1999 data and the FH8 alpha calculated using 2000 and 2001 data. Next, we predict the two-year-ahead FH8 alpha for year 2002 and 2003 at the beginning of year 2002 by applying the fitted model from training dataset on the testing dataset with 22 predictors. We repeat the procedure to predict two-year-ahead FH8 alpha for each fund that has 22 predictors available every year starting from 2002 to 2017.

2.2.4. Machine learning models

We use two linear models and three nonlinear regression tree-based models. The first linear model we consider is multiple linear regression model estimated via ordinary least squares (OLS). The second linear model we use is the least absolute shrinkage and selection operator (LASSO).[50] Unlike OLS, which estimates the coefficients for all the predictors, LASSO forces some of the irrelevant coefficient estimates to be zero and thus performs predictor selection.

Although linear models have advantages in interpretation, it is arbitrary to assume the relation between predictors and the dependent variable to be linear. Therefore, we also include

[50] See Tibshirani (1996) for more detail on the LASSO.

three nonlinear models in this study. There are many nonlinear machine learning models available. However, some of them are more suitable for classification problems such as support vector machine, and some of them perform more like a black box such as neural network. Considering the trade-off between flexibility and interpretation, we choose three tree-based models—bagged regression trees (Bag.Tree), random forest (RF), and boosted regression trees (Boost.Tree)—as the representatives of nonlinear models in this study.[51]

2.2.5. Prediction evaluation

The main research question of this study is whether machine learn can ex ante identify funds that deliver superior performance subsequently. To address this question, we train machine learning models to predict two-year-ahead FH8 alpha. We use two approaches to evaluate the predictions. The first approach uses statistical measures to compare the predicted values and the realized values in the testing dataset. First, we compute mean squared error (MSE), correlation coefficient, and out-of-sample R^2 between predicted and realized values of two-year-ahead FH8 alpha for all the five models. The out-of-sample R^2 (R^2_{OOS}) is calculated as:

$$R^2_{OOS} = 1 - \frac{\sum_1^n (y_i - \widehat{y_i})^2}{\sum_1^n (y_i - \overline{y_i})^2} \tag{2.1}$$

where y_i is the realized value of two-year-ahead FH8 alpha, $\widehat{y_i}$ is predicted value from a given machine learning model, and $\overline{y_i}$ is the mean of realized value of two-year-ahead FH8 alphas.

Next, to compare the prediction accuracy between two models we conduct the Diebold and Mariano (1995) test. The test statistic for comparing the predictions of model (1) and (2) is calculated as:

[51] See James, Witten, Hastie, and Tibshirani (2013) or Friedman, Hastie, and Tibshirani (2001) for detailed description of regression trees.

$$DM_{12} = \frac{\overline{d_{12}}}{\widehat{\sigma_{d_{12}}}} \tag{2.2}$$

where,

$$d_{12} = \frac{1}{m}\sum_{1}^{m}\left(\left(\hat{e}_j^{(1)}\right)^2 - \left(\hat{e}_j^{(2)}\right)^2\right) \tag{2.3}$$

$\hat{e}_j^{(1)}$ and $\hat{e}_j^{(2)}$ are the prediction error for fund j in a testing dataset using model (1) and model (2). Second and more importantly, we form portfolios on the predictions and then evaluate the out-of-sample performance of all the machine learning portfolios. This approach evaluates whether machine learning predictions add value to hedge fund investors.

2.3. Prediction accuracy evaluation

Table 2.2 reports the MSE, correlation and R^2_{OOS} between the predicted and realized FH8 alpha. In the testing dataset, not all funds that have 22 predictors exist in the next two years, therefore not all funds in the testing set have a realized FH8 alpha. However, since this is the case for all the models, the comparison is still a fair comparison. Table 2 shows that random forest performs the best, having lowest MSE, highest correlation and highest R^2_{OOS} among the five models, while OLS performs the worst, having the highest MSE, lowest correlation and lowest R^2_{OOS}.

Next, we conduct pairwise comparisons between models using Diebold and Mariano (1995) test. In Table 2.3, a positive value in the table suggests the model in the column has better prediction accuracy than the model in the row. Again, we find random forest outperforms any other four models and OLS underperforms any other four models although not significantly so. These initial assessments using statistical measures suggest that on average random forest shows great potential in predicting hedge fund performance.

2.4. Performance of machine learning portfolios

Above tests examine the prediction accuracy of all funds that have a realized two-year-ahead FH8 alpha. However, the more important question is whether machine learning predictions can identify the funds that deliver superior performance subsequently and thus create value to hedge fund investors. To answer this question, we form equal weight decile portfolios based on machine learning predictions and evaluate the out-of-sample performance of those portfolios. Specifically, at the beginning of each year starting from 2002 to 2017, we sort funds based on the predicted FH8 alpha from a given machine learning model into deciles. Considering the lockup and redemption restrictions hedge fund investors may face, we rebalance all the portfolios annually instead of commonly used monthly rebalancing in literature. For the funds stop reporting in the database in the holding period, we equally distribute the cumulative value of those funds to the remaining funds in the portfolio. This yields five top machine learning portfolios and five bottom machine learning portfolios from five models based on model predictions of FH8 alpha.

Figure 2.1 shows the value of one-dollar investment in each portfolio over the period of January 2002 to December 2017. For illustration purpose, we equal weight the top portfolios of OLS and LASSO and equal weight the top portfolios of the three tree-based models. For comparison, we also form an equal weight portfolio of all the hedge funds in the database and add cumulative return of SP500 over the same period. Figure 1 shows that one dollar invested in the top portfolio of linear machine learning models at the beginning of 2002 is worth \$7.03 at the end of 2017 and one dollar invested in the top portfolio of three based models is worth \$5.21. Meanwhile, the equal weight portfolio of all hedge funds only has a value of \$2.5, which is even lower than the value of SP500 for the same period. Bollen, Joenväärä, and Kauppila (2020) also confirm that hedge fund performance has weakened after 2008. Therefore, identifying winning funds ex ante becomes even more valuable.

Table 2.4 evaluates risk adjusted performance using FH8 model and other common performance metrics for all the machine learning portfolios. Panel A of Table 2.4 shows that all the top machine learning portfolios deliver positive and significant FH8 alpha. LASSO portfolio generates monthly FH8 alpha of 0.65% (*t*-stat = 3.52), and even the lowest top machine learning portfolio from OLS model delivers monthly FH8 alpha of 0.43% (*t*-stat = 2.24). Untabulated test of pairwise comparisons of the alphas from five top machine learning portfolios shows that LASSO statistically significantly outperforms other models.[52] The performance of all the bottom machine learning portfolios does not significantly differ from zero. Since we require funds to have past four years' returns to enter in the sample, all the sample funds have survived at least six years. Given the failure rate in hedge fund industry, the average sample fund is better than the average hedge fund of the industry. Therefore, it is not surprising that the bottom portfolios do not have significant negative alphas. Since investors cannot short hedge funds, we focus on top portfolios. Even so, the long short portfolios still have impressive positive alphas. Since we train all the models to predict FH8 alpha, we focus on evaluating whether all the portfolios deliver FH8 alpha out-of-sample. If we evaluate them with alternative factor models, all the machine learning portfolios still generate positive alphas, albeit statistically weaker.[53]

Panel B of Table 2.4 evaluates all the machine learning portfolios with model free measures such as average return, Sharpe ratio and attrition rate. Again, top LASSO portfolio has the highest

[52] The monthly alpha of the top portfolio from LASSO is 0.22% (*t*-stat = 2.21) higher than the one from OLS, 0.21% (*t*-stat = 2.25) higher than the one from boosted tree, 0.19% (*t*-stat = 1.92) higher than the one from bagged tree, and 0.16% (*t*-stat = 1.76) higher than the one from random forest.
[53] The results using Global 4 factor model and Global 7 factor model in Joenväärä, Kauppila, Kosowski, and Tolonen (2019) are similar.

average monthly return of 1.1%, the highest Sharpe ratio of 0.3 and lowest average annual attrition rate of 0.08 among the five top machine learning portfolios.

Although both press and research studies find the performance of hedge fund industry has declined in the recent decade, we investigate whether machine learning can still select a subset of funds that deliver superior performance for this period. To do so, we replicate our analysis in Table 2.4 for two subperiods. Motivated by Bollen, Joenväärä, and Kauppila (2020), which finds 2008 is the crucial point of the change of the performance of hedge fund industry, we also use 2008 as the break point. Therefore, we evaluate the all the machine learning portfolios over the two subperiods: January 2002 to December 2007 and January 2008 to December 2017. Table 2.5 presents the results. All the top machine learning portfolios produce positive FH8 alphas over the two subperiods. FH8 alphas of five top portfolios are weaker during 2002 to 2007 but even stronger during 2008 to 2017. Fewer observations in the first subperiod may partially explain the lower *t*-statistics for this period. However, LASSO consistently produce the highest FH8 alpha among the five models over each period.

In terms of other performance metrics, although the average return and Sharpe ratio of machine learning portfolios declined over 2008 to 2017, they still deliver impressive returns. For example, the LASSO portfolio has an average monthly return of 0.97%, a Sharpe ratio of 0.23, while SP500 has an average monthly return of 0.75% and a Sharpe ratio of 0.17 over the same period.

Taken together, the results from Table 2.4 to Table 2.5 show that all the machine learning portfolios are able to select a subset of funds that outperform subsequently, even for the period over 2008 to 2017 during which the performance of hedge fund industry on average is disappointing. Among the five machine learning portfolios, LASSO consistently outperform other models in identifying winning funds.

2.5. Predictor importance

2.5.1. Average relative importance

We next examine which predictors are more important among all the 22 predictors. Figure 2.2 reports the most important ten predictors in OLS, LASSO, random forest and boosted regression trees.[54] The predictor importance of OLS is ranked using the average of absolute value of *t*-statistic of each coefficient across 16 sample years. The predictor importance of LASSO is ranked using the average of absolute value of *t*-statistic of coefficient estimates from OLS estimation of the selected predictors by LASSO. Predictor importance of random forest is calculated using the mean decrease in Gini index. Importance of boosted regression trees is calculated from the relative influence of predictors. All the measures are scaled so the most important predictor has a value of 100.

It is important to note that all the rankings are within a given model. However, we still find that several predictors are ranked in top ten by all models. They are 1-year average return, 4-year average return, MAX, 2-year alpha, 3-year alpha, 2-year systematic risk, 4-year systematic risk and beta activity. In contrast, 2-year R^2, 3-year R^2, 4-year R^2, 2-year Talpha, 3-year Talpha, 4-year Talpha, 3-year Sharpe ratio and 4-year Sharpe ratio are ranked in the bottom ten by all models.

Overall, average return, MAX, FH8 alpha, systematic risk, and beta activity are consistently selected as important predictors for FH8 alpha by different machine learning models, while R^2, Talpha and Sharpe ratio are relatively not so important. In addition, these results also

[54] Since bagged regression trees do not provide variable importance, we only report the relative importance for the other four models.

suggest that predictors with different horizons each provide incremental information on predicting fund future performance.

2.5.2. Time-varying importance

Prior work on individual performance predictors usually implicitly assumes the predictor is of the same importance across time. However, Sun, Wang, and Zheng (2018) show that the persistence of hedge fund performance only shows up following weak hedge fund market conditions but not in strong markets. Kacperczyk, van Nieuwerburgh and Veldkamp (2014) also show that skilled mutual managers focus more on stock picking during boom market while focus more on market timing during recession.

Since fund managers shift their skill with different market states and fund performance persistence is conditional on market conditions, it is suboptimal to assume the importance of a predictor to be constant. Using machine learning to predict subsequent fund performance, we relax this assumption and allow the importance of a predictor to vary over time. The empirical results also show that the importance of predictors indeed varies considerably over time.

Figure 2.3 shows the relative importance of past 2-year alpha and beta activity from LASSO during our sample period. A value of 100 means the predictor is selected by LASSO as the most important predictor among all the 22 predictors in that year. A value of zero means the predictor is not selected by LASSO as a relevant predictor in that year. For example, LASSO identifies beta activity as the most important predictor in 2002, 2004, 2006 and 2008 and suggests past 2-year alpha is not so important comparted to other predictors in 2002, 2004, 2006 but is the most important predictor in 2017. Figure 4 shows that absolute value of *t*-statistic from OLS estimation of a predictor that is selected by LASSO. Figure 2.4 further confirms that there is a large variation on the importance of 2-year alpha and beta activity over time in absolute terms.

For illustration purpose, we only present the time-varying importance of 2-year alpha and beta activity. However, the conclusion holds for all the predictors in this study. Consistent with the findings of prior literature, the importance of hedge fund predictors varies substantially over time.

2.6. Performance of portfolios formed on individual predictors

In this section, we examine the performance of the portfolios formed on individual predictors over the same period as the machine learning portfolios. Using the same set of sample funds, we form equal weight decile portfolios on each individual predictor and rebalance the portfolios annually. Table 2.6 reports the FH8 alphas of the individual portfolios.

Table 2.6 shows that FH8 alpha differs dramatically across different predictors. Beta activity generates 0.5% (*t*-stat = 3.33) monthly alpha, which is the highest among the 22 individual predictors. However, alpha of top portfolio formed on past 3-year average return and alpha of top portfolio formed on past 4-year average return do not significantly differ from zero. Contrastingly, all the top machine learning portfolios in Table 2.4 delivers a significant positive alpha. Even the lowest alpha from machine learning portfolio has 0.43% per month. The average alpha of all the top machine learning portfolios is 0.49% per month, while the average alpha of all the top individual portfolios is only 0.31% per month. These results show the advantage of using all the individual predictors with machine learning than using individual predictors.

We also examine the performance of portfolios formed on individual predictors using model free measures. Table 2.7 reports the average monthly return, Sharpe ratio, and average annual attrition rate of all the individual portfolios. The performance of individual portfolios tends to align with the properties of the individual predictors. For example, portfolios formed on average return or systematic risk tend to have higher out-of-sample average return, and portfolios formed on Sharpe ratio tend to have higher out-of-sample Sharpe ratio. Both 3-year systematic risk and 4-year systematic risk generate highest average return, which is 0.94% per month. However, this is still lower than the average return of machine learning portfolio from LASSO, which generates an average return of 1.10% per month. To sum up, none of the portfolios formed on individual predictors has higher out-of-sample alpha than the machine learning portfolio using LASSO. Also, machine learning portfolios have good properties in all the common performance metrics while the performance of portfolios formed on individual predictors tend to align with the property of the given individual predictor.

2.7. Machine learning portfolios and individual portfolios

To better understand the contribution of machine learning predictions and individual predictors, we isolate the effects from machine learning predictions and individual predictors by evaluating the performance of mutually exclusive portfolios. Specifically, at the beginning of each year starting from 2002 to 2017, we sort funds into quintiles on two measures separately. Then we form one mutually exclusive portfolio with funds in the top quintile of measure A but not in the top quintile of measure B and form the other mutually exclusive portfolio with funds in the top quintile of measure B but not in the top quintile of measure A.

We use the machine learning portfolios from LASSO and OLS—the strongest and the weakest machine learning portfolios in Table 2.4—to from mutually exclusive portfolios with classical individual predictors in literature and with individual predictors that demonstrate predictive power in our previous analysis.

Table 2.8 shows that performance of mutually exclusive portfolios formed on LASSO prediction and individual predictors. Let us take the performance of top LASSO and 1-R2_2year portfolios after excluding the funds that are common to them for example. After we exclude funds in the top 1-R2_2year quintile from the top LASSO predictions quintile, the top LASSO portfolio still delivers a significant alpha, while the top 1-R2_2year portfolio excluding funds common to the top LASSO portfolio barely retain the significance in alpha. The same pattern presents in all the mutually exclusive portfolios. In all the cases, LASSO portfolios retain its significant alpha after excluding the common funds to the top portfolio of individual predictor, while in several cases the alpha of individual portfolios totally disappeared after excluding the funds in the top LASSO portfolio. For example, beta activity and MAX generate significant alpha in Table 2.6 (0.5% and 0.4% separately), however, once we isolate the effect from LASSO prediction, those two predictors deliver no significant alpha out-of-sample.

In addition, after excluding the common funds, LASSO portfolio almost always dominates the portfolio of an individual predictor in all other performance metrics, delivering a higher average return, higher Sharpe ratio, and lower attrition rate. Except for a few cases, the LASSO portfolio has a slightly higher attrition rate. The last column of Table 2.8 reports the proportion of funds in one top portfolio but not in the other top portfolio. This measure shows that there is a substantial proportion of funds identified by the LASSO prediction as outperforming funds but not by individual predictors.

We further evaluate the performance of machine learning portfolios and individual portfolios by forming the mutually exclusive portfolios using the weakest machine learning predictions from OLS with individual predictors in Table 2.9. Table 2.9 shows a similar pattern as Table 2.8. Machine learning portfolio still delivers significant and positive alpha after we isolate the effect from individual predictors. However, the performance of individual predictors deteriorates once we isolate the contribution of machine learning predictions. For example, the alphas of 2-year systematic risk and MAX disappear after we exclude the top funds identified by machine learning.

These results show that the machine learning portfolios retain their long-term out-of-sample predictive power while the reverse is not true for individual predictors, suggesting machine learning methods unveil unique information of future hedge fund performance that is not captured by individual predictors.

2.8. Conclusion

Although the performance of the hedge fund industry has, on average, been disappointing in the past decade, we find that machine learning models can still consistently identify funds that outperform. We train five machine learning models to predict two-year-ahead alpha using 22 measures. Based on those predictions, we then form equal weight, annually rebalanced portfolios. The machine learning portfolio formed using LASSO delivers, over the period 2002 to 2017, a 7.8% alpha per year, a 13.2% average return per year, and a 0.3 Sharpe ratio.

Unlike traditional methods, which typically sort on one predictor and implicitly assume time-invariant predictive power, we train our machine learning models to continually re-identify the important predictors. We find that average return, alpha, MAX, systematic risk, and beta activity are consistently selected as important predictors, but the importance of all of the predictors is time-varying. Since Harvey, Liu, and Zhu (2016), there is a growing debate about the persistence of predictors in the literature. Our models acknowledge that debate by dynamically updating based on new information about the utility of each predictor.

To compare with traditional methods, we also form 22 portfolios based on the individual predictors. We find that, in terms of alpha, none of the individual predictor portfolios outperforms the machine learning portfolio formed using LASSO. The average alpha of the individual predictor portfolios is 3.72% per year, compared to 5.88% per year for the machine learning portfolios. Isolating the contribution of the individual predictors shows that machine learning process provides unique information not captured by the individual predictors.

Table 2.1: Summary statistics: 1994 – 2018

This table reports the summary statistics for all hedge funds between January 1994 and December 2018. Panel A reports the cross-sectional mean, median, 10^{th} percentile, 90^{th} percentile, standard deviation of fund characteristics. Panel B reports the number of hedge funds and the average and standard deviation of monthly returns of an equal weight portfolio of all hedge funds in the database by year.

Panel A. Fund characteristics

	Mean	Median	P10	P90	Std.Dev.
Monthly excess return (%)	0.35	0.39	-4.08	4.73	3.85
Assets (millions $)	158.85	45.40	3.97	428.15	305.65
Minimum investment (million $)	0.85	0.25	0.03	1.00	1.71
Management fee (%)	1.46	1.50	0.90	2.00	0.57
Performance fee (%)	17.20	20.00	5.00	20.00	6.59
Hurdle rate (%)	0.31	0.00	0.00	0.00	1.43
Lockup period (days)	61.12	0.00	0.00	360.00	145.86
Redemption notice (days)	15.25	0.00	0.00	60.00	24.74
Longevity (months)	109.03	97.00	32.00	207.00	67.41

Table 2.1 (Cont.)

Panel B: Equal-weight portfolio of all hedge funds

Year	Number of funds	Mean Return	Std.Dev.
1994	309	0.19	0.96
1995	402	1.77	1.27
1996	561	1.68	1.65
1997	741	1.54	1.99
1998	922	0.59	2.21
1999	1318	2.08	2.28
2000	1726	0.85	2.42
2001	2322	0.51	1.28
2002	3145	0.55	1.20
2003	4107	1.83	1.40
2004	5163	0.88	1.60
2005	6204	0.52	1.44
2006	7100	1.21	1.74
2007	7940	1.08	1.71
2008	8213	-1.61	3.59
2009	7737	1.66	2.59
2010	7816	0.76	2.80
2011	7876	-0.50	2.56
2012	7361	0.47	1.80
2013	6837	0.63	1.38
2014	6382	-0.03	1.21
2015	6024	-0.23	1.36
2016	5549	0.19	1.58
2017	5205	1.00	0.59
2018	4811	-0.69	1.75

Table 2.2: Prediction accuracy

This table shows the mean squared error, correlation and R^2_{OOS} of predicted value and realized value of two-year-ahead FH8 alpha for five models.

	OLS	LASSO	Bag.Tree	RF	Boost.Tree
MSE	3.11	1.45	1.43	1.39	1.65
Correlation	0.07	0.11	0.15	0.15	0.10
R^2_{OOS}	-1.59	-0.21	-0.19	-0.16	-0.37

Table 2.3: Comparing predictions using Diebold-Mariano tests

This table shows pairwise Diebold-Mariano test statistics comparing prediction accuracy among five models. A positive number indicates the model in the column outperforms the model in the row. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

	LASSO	Bag.Tree	RF	Boost.Tree
OLS	1.45	1.45	1.48	1.31
LASSO		-0.20	0.49	-3.27***
Bag.Tree			3.32***	-2.92***
RF				-3.98***

Table 2.4: Performance of machine learning portfolios

This table shows the performance of annually rebalanced portfolios of hedge funds formed based on machine learning predictions. At the beginning of each year, we sort the funds into deciles based on their predicted two-year-ahead FH8 alpha from a machine learning model. The 'Top' portfolio is an equal weight portfolio of all the funds in the highest predicted value decile. The 'Btm' portfolio is an equal weight portfolio of all the funds in the lowest predicted value decile. The period of evaluation for the portfolios is January 2002 through December 2017. Panel A reports the FH8 alpha of all the portfolios. *t*-statistics are in brackets below each alpha ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels. Panel B shows the average return, Sharpe ratio, and average annual attrition rate of all the portfolios.

Panel A: FH8 alpha

	OLS	LASSO	Bag.Tree	RF	Boost.Tree
Top	0.43**	0.65***	0.46***	0.49***	0.44***
	[2.24]	[3.52]	[2.84]	[2.90]	[2.92]
Btm	0.01	-0.09	-0.11	-0.16	-0.00
	[0.09]	[-0.59]	[-0.85]	[-1.16]	[-0.02]
Top - Btm	0.42	0.75**	0.57**	0.65***	0.44**
	[1.48]	[2.58]	[2.54]	[2.65]	[2.19]

Panel B: Other performance measures

		OLS	LASSO	Bag.Tree	RF	Boost.Tree
	Top	0.87	1.10	0.80	0.84	0.80
Average Return	Btm	0.31	0.23	0.22	0.19	0.32
	Top - Btm	0.56	0.87	0.58	0.64	0.48
	Top	0.24	0.30	0.25	0.26	0.25
Sharpe Ratio	Btm	0.12	0.08	0.08	0.07	0.13
	Top - Btm	0.12	0.23	0.18	0.19	0.12
	Top	0.10	0.08	0.11	0.10	0.10
Attrition Rate	Btm	0.12	0.14	0.12	0.13	0.12
	Top - Btm	-0.03	-0.06	-0.02	-0.03	-0.02

Table 2.5: Performance of machine learning portfolios – subperiods

This table shows the results for two subperiods of the tests in Table 2.4. Panel A reports the FH8 alpha of all the machine learning portfolios in Table 2.4 for the period of January 2002 to December 2007. Panel B reports the FH8 alpha of all the machine learning portfolios in Table 2.4 for the period of January 2008 to December 2017. Panel C reports the average return, Sharpe ratio, and average annual attrition rate of all the portfolios for the period of January 2002 to December 2007. Panel D reports the average return, Sharpe ratio, and average annual attrition rate of all the portfolios for the period of January 2008 to December 2017.

Panel A: FH8 alpha (2002 – 2007)

	OLS	LASSO	Bag.Tree	RF	Boost.Tree
Top	0.38*	0.72***	0.25	0.31	0.39*
	[1.97]	[2.81]	[1.13]	[1.28]	[1.75]
Btm	0.26	0.20	0.29	0.29	0.34**
	[1.33]	[0.82]	[1.61]	[1.57]	[2.09]
Top - Btm	0.12	0.52	-0.04	0.02	0.05
	[0.45]	[1.30]	[-0.13]	[0.05]	[0.19]

Panel B: FH8 alpha (2008 – 2017)

	OLS	LASSO	Bag.Tree	RF	Boost.Tree
Top	0.71**	0.74***	0.62***	0.62**	0.52**
	[2.50]	[2.91]	[2.69]	[2.60]	[2.50]
Btm	-0.27	-0.35*	-0.28	-0.36*	-0.22
	[-1.51]	[-1.77]	[-1.56]	[-1.92]	[-1.27]
Top - Btm	0.98**	1.09***	0.90***	0.98***	0.74**
	[2.47]	[2.77]	[2.85]	[2.81]	[2.60]

Panel C: Other performance measures (2002 – 2007)

		OLS	LASSO	Bag.Tree	RF	Boost.Tree
	Top	0.83	1.32	0.96	1.00	1.06
Average Return	Btm	1.08	0.99	0.83	0.82	0.99
	Top - Btm	-0.26	0.34	0.13	0.18	0.07
	Top	0.36	0.52	0.38	0.39	0.41
Sharpe Ratio	Btm	0.41	0.34	0.38	0.40	0.47
	Top - Btm	-0.05	0.18	0.00	0.00	-0.06
	Top	0.10	0.08	0.11	0.10	0.10
Attrition Rate	Btm	0.12	0.14	0.12	0.13	0.12
	Top - Btm	-0.03	-0.06	-0.02	-0.03	-0.02

Panel D: Other performance measures (2008 – 2017)

		OLS	LASSO	Bag.Tree	RF	Boost.Tree
	Top	0.89	0.97	0.70	0.74	0.65
Average Return	Btm	-0.16	-0.23	-0.15	-0.18	-0.08
	Top - Btm	1.05	1.20	0.85	0.92	0.73
	Top	0.21	0.23	0.20	0.21	0.19
Sharpe Ratio	Btm	-0.06	-0.08	-0.05	-0.06	-0.03
	Top - Btm	0.27	0.32	0.25	0.27	0.22
	Top	0.10	0.08	0.11	0.10	0.10
Attrition Rate	Btm	0.12	0.14	0.12	0.13	0.12
	Top - Btm	-0.03	-0.06	-0.02	-0.03	-0.02

Table 2.6: Performance of portfolios formed on individual predictors – FH8 alpha

This table shows the FH8 alpha of annually rebalanced portfolios of hedge funds formed on one predictor of all the 22 predictors in this study. At the beginning of each year, we sort the funds into deciles based their value of a given predictor. The 'Top' portfolio is an equal weight portfolio of all the funds in the highest value of the predictor. The 'Btm' portfolio is an equal weight portfolio of all the funds in the lowest value of the predictor. The period of evaluation for the portfolios is January 2002 through December 2017. *t*-statistics are in brackets below each alpha ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Ret_1year	Sharpe_1year	MAX
Top	0.32*	0.33***	0.40**
	[1.79]	[4.35]	[2.24]
Btm	0.21	0.05	0.09*
	[1.23]	[0.41]	[1.71]
Top - Btm	0.12	0.28**	0.31*
	[0.41]	[2.07]	[1.73]

	Ret_2year	Sharpe_2year	Alpha_2year	Talpha_2year	1-R2_2year	Sys.Risk_2year
Top	0.29*	0.29***	0.39***	0.35***	0.19***	0.31*
	[1.70]	[4.23]	[3.02]	[6.35]	[2.69]	[1.78]
Btm	0.22	0.08	0.14	0.06	0.09	0.18***
	[1.35]	[0.65]	[1.00]	[0.57]	[0.94]	[4.80]
Top - Btm	0.06	0.21	0.25	0.29***	0.11	0.13
	[0.24]	[1.55]	[1.45]	[2.72]	[1.06]	[0.76]

	Ret_3year	Sharpe_3year	Alpha_3year	Talpha_3year	1-R2_3year	Sys.Risk_3year
Top	0.20	0.24***	0.43***	0.30***	0.19***	0.31*
	[1.25]	[3.75]	[3.26]	[6.06]	[2.84]	[1.77]
Btm	0.32**	0.20	0.10	0.09	0.02	0.19***
	[2.00]	[1.48]	[0.76]	[0.78]	[0.22]	[5.03]
Top - Btm	-0.12	0.04	0.33*	0.21*	0.17*	0.12
	[-0.48]	[0.30]	[1.95]	[1.97]	[1.86]	[0.67]

	Ret_4year	Sharpe_4year	Alpha_4year	Talpha_4year	1-R2_4year	Sys.Risk_4year	Beta Activity
Top	0.25	0.22***	0.44***	0.28***	0.21***	0.34**	0.50***
	[1.58]	[3.63]	[3.49]	[5.75]	[3.36]	[1.99]	[3.33]
Btm	0.14	0.09	0.01	0.01	0.04	0.20***	0.06
	[0.94]	[0.69]	[0.10]	[0.06]	[0.53]	[5.17]	[0.41]
Top - Btm	0.11	0.13	0.43***	0.27**	0.17*	0.14	0.44**
	[0.46]	[0.88]	[2.64]	[2.60]	[1.88]	[0.80]	[2.18]

Table 2.7: Performance of portfolios formed on individual predictors – other measures

This table shows the average return, Sharpe ratio, and average annual attrition rate of all the portfolios in Table 2.6.

		Ret_1year	Sharpe_1year	MAX
Average Return	Top	0.75	0.47	0.92
	Btm	0.56	0.26	0.17
	Top - Btm	0.19	0.21	0.75
Sharpe Ratio	Top	0.24	0.40	0.25
	Btm	0.16	0.12	0.19
	Top - Btm	0.08	0.28	0.06
Attrition Rate	Top	0.06	0.06	0.10
	Btm	0.17	0.18	0.14
	Top - Btm	-0.11	-0.12	-0.05

		Ret_2year	Sharpe_2year	Alpha_2year	Talpha_2year	1-R2_2year	Sys.Risk_2year
Average Return	Top	0.64	0.39	0.68	0.49	0.35	0.93
	Btm	0.59	0.33	0.60	0.44	0.63	0.25
	Top - Btm	0.05	0.06	0.08	0.05	-0.28	0.68
Sharpe Ratio	Top	0.20	0.34	0.27	0.42	0.29	0.21
	Btm	0.19	0.15	0.19	0.17	0.18	0.38
	Top - Btm	0.01	0.19	0.09	0.25	0.12	-0.17
Attrition Rate	Top	0.06	0.06	0.07	0.07	0.12	0.11
	Btm	0.18	0.19	0.16	0.17	0.09	0.11
	Top - Btm	-0.12	-0.13	-0.09	-0.10	0.04	-0.01

Table 2.7 (Cont.)

		Ret_3year	Sharpe_3year	Alpha_3year	Talpha_3year	1-R2_3year	Sys.Risk_3year
Average Return	Top	0.60	0.35	0.72	0.42	0.32	0.94
	Btm	0.64	0.47	0.54	0.46	0.58	0.25
	Top - Btm	-0.04	-0.11	0.18	-0.04	-0.26	0.68
Sharpe Ratio	Top	0.19	0.32	0.31	0.41	0.30	0.21
	Btm	0.22	0.20	0.17	0.18	0.16	0.39
	Top - Btm	-0.03	0.12	0.14	0.24	0.14	-0.19
Attrition Rate	Top	0.05	0.07	0.06	0.07	0.13	0.11
	Btm	0.19	0.20	0.15	0.17	0.09	0.12
	Top - Btm	-0.13	-0.14	-0.09	-0.10	0.04	-0.01

		Ret_4year	Sharpe_4year	Alpha_4year	Talpha_4year	1-R2_4year	Sys.Risk_4year	Beta Activity
Average Return	Top	0.62	0.33	0.71	0.40	0.34	0.94	0.83
	Btm	0.55	0.42	0.49	0.40	0.63	0.27	0.47
	Top - Btm	0.07	-0.09	0.23	0.00	-0.29	0.67	0.36
Sharpe Ratio	Top	0.19	0.30	0.30	0.38	0.34	0.21	0.32
	Btm	0.19	0.18	0.15	0.15	0.17	0.41	0.14
	Top - Btm	0.00	0.13	0.15	0.23	0.18	-0.20	0.18
Attrition Rate	Top	0.05	0.06	0.06	0.06	0.13	0.10	0.11
	Btm	0.20	0.21	0.19	0.19	0.08	0.11	0.09
	Top - Btm	-0.15	-0.15	-0.12	-0.13	0.05	-0.01	0.01

Table 2.8: Machine learning portfolio (LASSO) and classical portfolio

This table shows the results of annually rebalanced portfolios double sorted on machine learning prediction from LASSO and classical individual predictors. For example, LASSO excl. 1-R2_2year is the portfolio of funds in the top quintile of machine learning prediction from LASSO, excluding funds in the top quintile of 1-R2_2year measure. The period of evaluation for the portfolios is January 2002 through December 2017. *t*-statistics are in brackets below FH8 alpha ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Attrition Rate is the average annual rate at which funds stop reporting. Proportion is the average of the proportion of funds in the top quintile of one measure but not in the top quintile of another measure.

	FH8 alpha	Average Return	Sharpe Ratio	Attrition Rate	Proportion
LASSO excl. 1-R2_2year	0.55***	0.97	0.31	0.07	0.75
	[3.57]				
1-R2_2year excl. LASSO	0.15*	0.30	0.23	0.13	
	[1.86]				
LASSO excl. Sys.Risk_3year	0.46***	0.70	0.39	0.08	0.65
	[4.51]				
Sys.Risk_3year excl. LASSO	0.07	0.53	0.15	0.10	
	[0.47]				
LASSO excl. Sys.Risk_2year	0.45***	0.69	0.37	0.08	0.66
	[4.55]				
Sys.Risk_2year excl. LASSO	0.05	0.49	0.14	0.10	
	[0.30]				
LASSO excl. Beta Activity	0.44***	0.84	0.26	0.08	0.76
	[2.62]				
Beta Activity excl. LASSO	0.16	0.45	0.22	0.13	
	[1.39]				
LASSO excl. MAX	0.46***	0.79	0.33	0.08	0.68
	[3.91]				
MAX excl. LASSO	0.11	0.52	0.17	0.09	
	[0.74]				
LASSO excl. Ret_1year	0.56***	0.88	0.31	0.09	0.67
	[3.72]				
Ret_1year excl. LASSO	0.17	0.54	0.20	0.07	
	[1.15]				
LASSO excl.Sharpe_1year	0.58***	0.97	0.31	0.09	0.72
	[3.58]				
Sharpe_1year excl. LASSO	0.21**	0.39	0.27	0.07	
	[2.51]				
LASSO excl. Alpha_2year	0.55***	0.95	0.32	0.09	0.68
	[3.50]				
Alpha_2year excl. LASSO	0.25**	0.51	0.23	0.07	
	[2.19]				
LASSO excl. Alpha_3year	0.53***	0.90	0.31	0.08	0.68
	[3.47]				
Alpha_3year excl. LASSO	0.23**	0.46	0.22	0.07	
	[2.04]				
LASSO excl. Talpha_3year	0.57***	0.97	0.31	0.09	0.73
	[3.50]				
Talpha_3year excl. LASSO	0.27***	0.43	0.30	0.07	
	[4.18]				

Table 2.9: Machine learning portfolio (OLS) and classical portfolio

This table shows the results of annually rebalanced portfolios double sorted on machine learning prediction from OLS and classical individual predictors. For example, OLS excl. 1-R2_2year is the portfolio of funds in the top quintile of machine learning prediction from OLS, excluding funds in the top quintile of 1-R2_2year measure. The period of evaluation for the portfolios is January 2002 through December 2017. *t*-statistics are in brackets below FH8 alpha ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels. Attrition Rate is the average annual rate at which funds stop reporting. Proportion is the average of the proportion of funds in the top quintile of one measure but not in the top quintile of another measure.

	FH8 alpha	Average Return	Sharpe Ratio	Attrition Rate	Proportion
OLS excl. 1-R2_2year	0.40**	0.78	0.25	0.08	
	[2.58]				0.78
1-R2_2year excl. OLS	0.20***	0.33	0.26	0.12	
	[2.69]				
OLS excl. Sys.Risk_3year	0.37***	0.60	0.34	0.09	
	[3.82]				0.64
Sys.Risk_3year excl. OLS	0.05	0.56	0.15	0.10	
	[0.28]				
OLS excl. Sys.Risk_2year	0.35***	0.58	0.31	0.09	
	[3.67]				0.65
Sys.Risk_2year excl. OLS	0.00	0.51	0.14	0.10	
	[0.02]				
OLS excl. Beta Activity	0.31*	0.66	0.22	0.08	
	[1.97]				0.73
Beta Activity excl. OLS	0.27**	0.56	0.27	0.11	
	[2.40]				
OLS excl. MAX	0.37***	0.66	0.28	0.09	
	[3.07]				0.66
MAX excl. OLS	0.12	0.57	0.18	0.09	
	[0.79]				
OLS excl. Ret_1year	0.42***	0.73	0.26	0.11	
	[2.85]				0.70
Ret_1year excl. OLS	0.19	0.58	0.21	0.06	
	[1.26]				
OLS excl. Sharpe_1year	0.40**	0.78	0.25	0.10	
	[2.51]				0.72
Sharpe_1year excl. OLS	0.25***	0.44	0.31	0.07	
	[3.10]				
OLS excl. Alpha_2year	0.36**	0.73	0.24	0.10	
	[2.31]				0.73
Alpha_2year excl. OLS	0.29**	0.53	0.25	0.06	
	[2.59]				
OLS excl. Alpha_3year	0.34**	0.71	0.24	0.10	
	[2.26]				0.74
Alpha_3year excl. OLS	0.23**	0.47	0.23	0.07	
	[2.13]				
OLS excl. Talpha_3year	0.38**	0.78	0.25	0.10	
	[2.41]				0.77
Talpha_3year excl. OLS	0.38**	0.46	0.31	0.06	
	[2.41]				

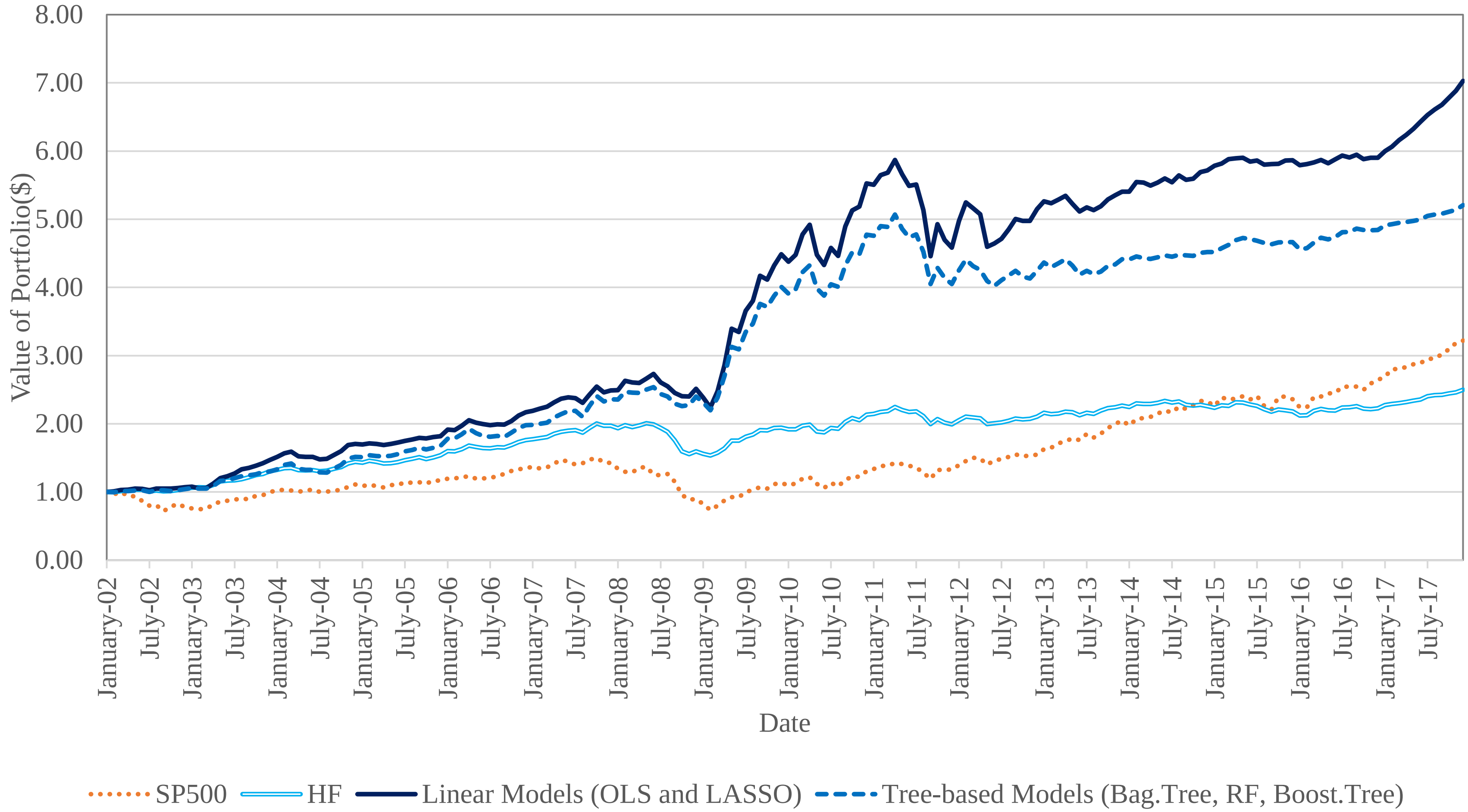

Figure 2.1: Value of machine learning portfolios

This figure shows the changing value of $1 invested in January 2002 through December 2017 in three equal weight portfolios of mutual funds and SP500 index. The HF portfolio is an equal weight portfolio of all hedge funds. Portfolio of linear models is an equal weight portfolio of two top decile portfolios formed on the predicted value of two-year-ahead FH8 from LASSO and OLS. Portfolio of tree-based models is an equal weight portfolio of three top decile portfolios formed on the predicted value of two-year-ahead FH8 from bagged trees, random forest and boost trees.

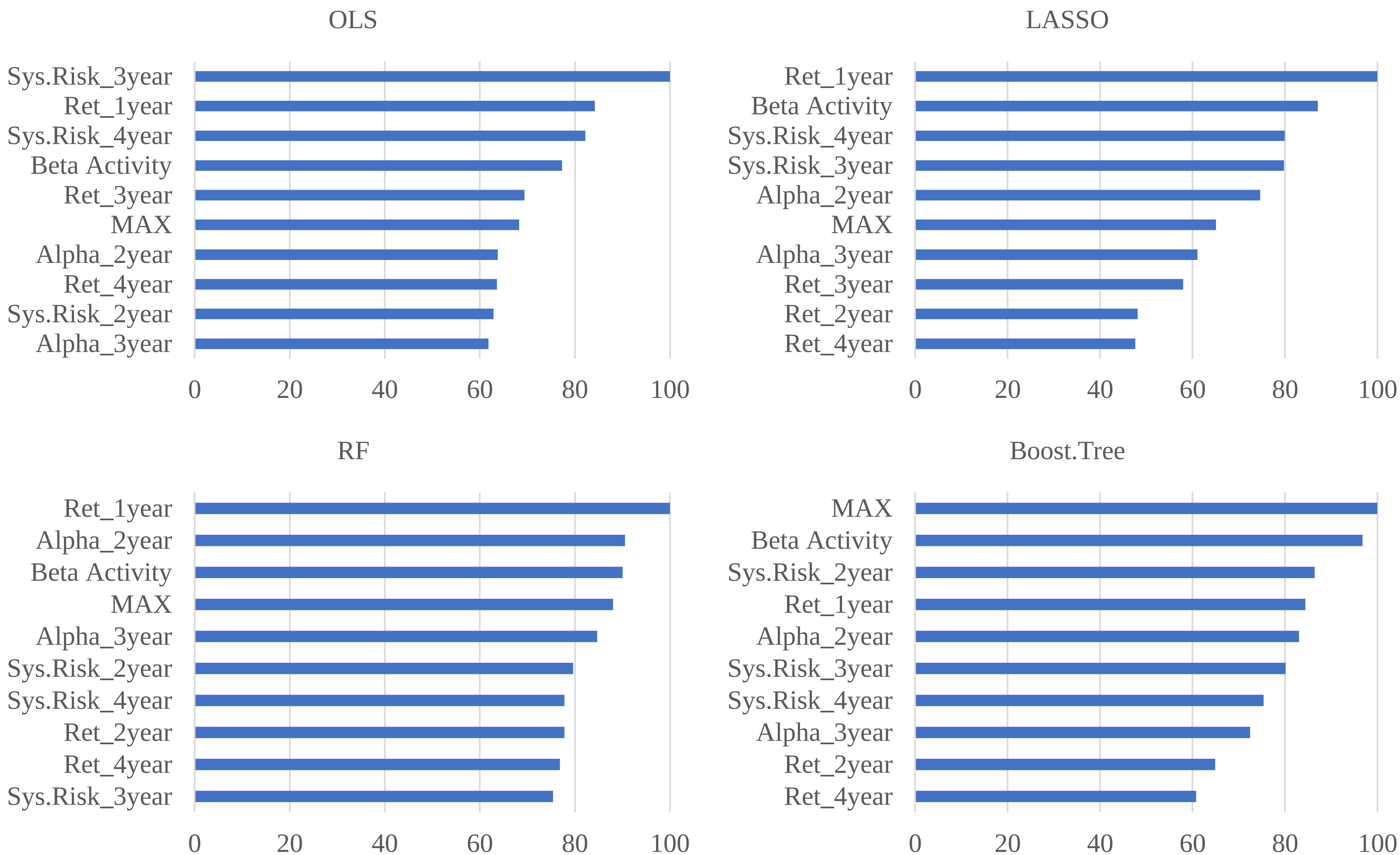

Figure 2.2: Predictor importance
This figure shows the top 10 most important predictors in each model. Predictor importance is an average of all training samples. The predictor with the highest importance is normalized to 100.

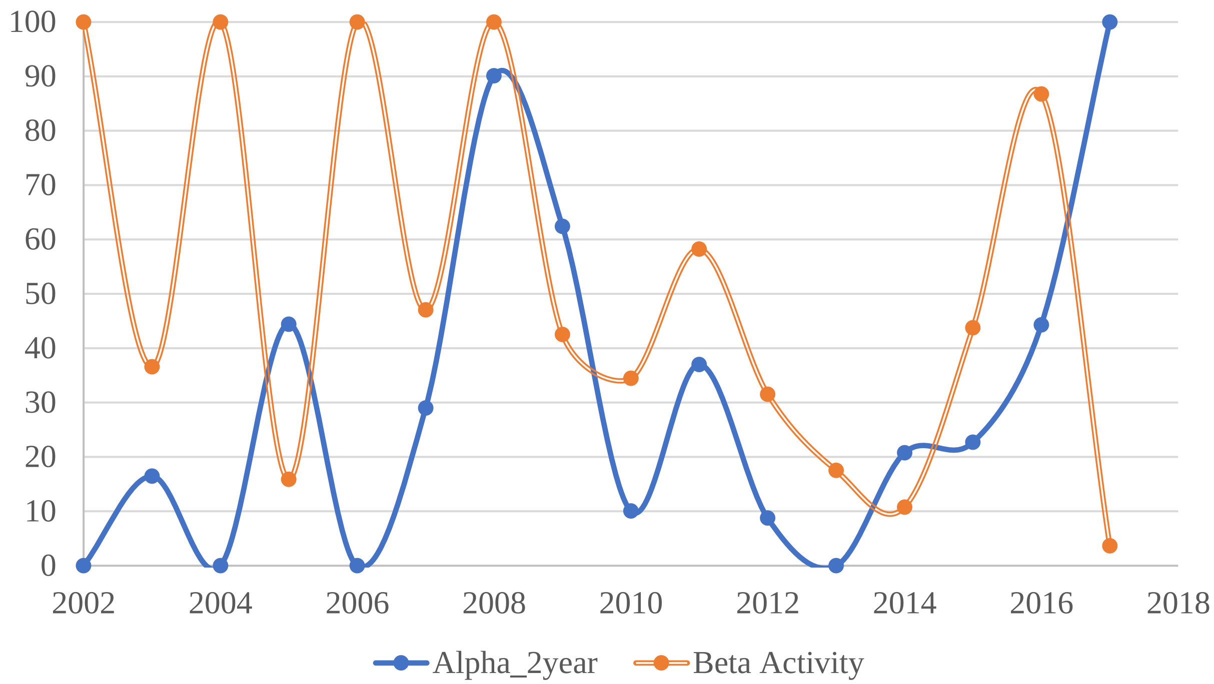

Figure 2.3: Predictor time vary importance – relative importance

This figure shows the relative importance of Alpha_2year and Beta Activity from LASSO each year from 2002 to 2017. A value of 100 indicates the preidctor is the most important predictor among 22 predictor in that year. A value of 0 indicates the preidctor is not selected by LASSO in that year.

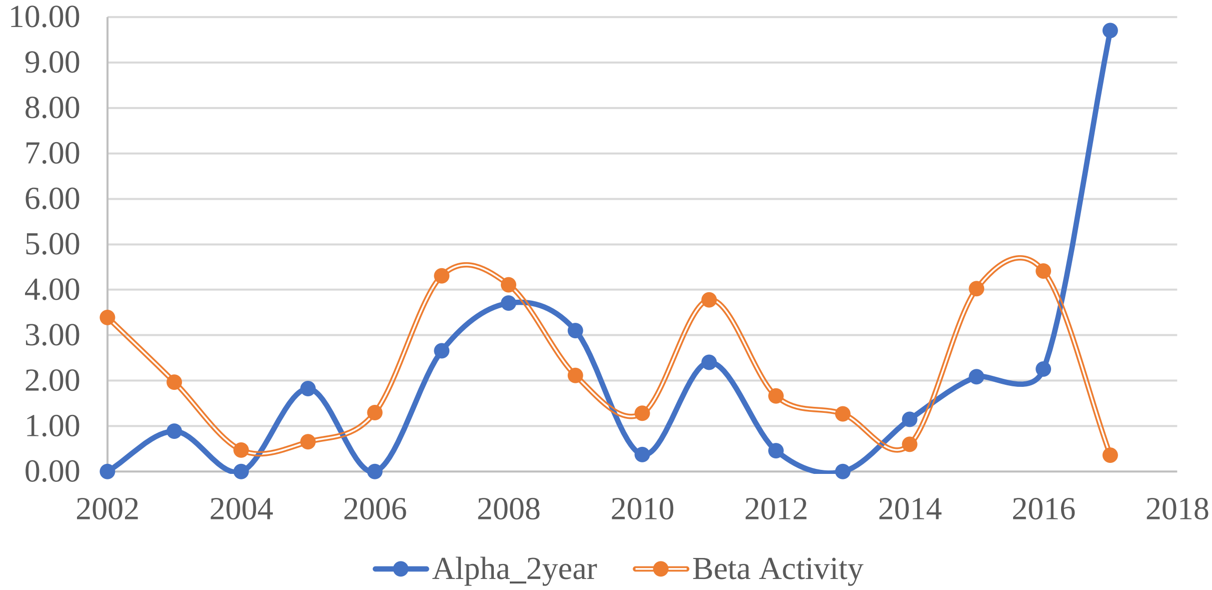

Figure 2.4: Predictor time vary importance – absolute importance
This figure shows the absolute importance of Alpha_2year and Beta Activity from LASSO each year from 2002 to 2017. The value of importance is the absolute value of *t*-statistic from OLS regression post LASSO selection.

Chapter 3. How Risky Is Your Manager? Maximum Drawdown as Predictor of Mutual Fund Performance and Flows

Timothy B. Riley, Qing Yan

3.1. Introduction

Academic studies of actively managed mutual funds tend to focus on risk-adjusted performance, not on risk alone. When risk is directly considered, it is typically measured as volatility or beta. However, in practice, maximum drawdown—which measures the largest decline in a fund's value from peak to trough—is widely used. In this study, we comprehensively evaluate this standard industry measure. We consider both (i) whether maximum drawdown contains useful information for investors and (ii) how investors approach maximum drawdown when making capital allocations. In the process, we find that a number of mutual fund managers have substantial risk management skills.

Maximum drawdown is not a complete measure of mutual fund risk, but for our purposes, it has some key characteristics. First, maximum drawdown focuses only on downside risk. A fund that continually increases in value by non-constant increments will have volatility, but no maximum drawdown. Second, maximum drawdown is easy for any investor to understand. Unlike other downside risk measures, such as semivariance, maximum drawdown can be quickly grasped by an investor with low mathematical and financial sophistication. Third, maximum drawdown is salient to investors. Morningstar's website reports maximum drawdown on the same tab as volatility and beta, and previous research indicates that the saliency of information has a large effect on whether investors consider the information (e.g., Barber, Odean, and Zheng, 2005).

We begin by estimating how risky actively managed U.S. equity mutual funds are in terms of maximum drawdown. Using 12-month periods from 1999 through 2018, the average maximum drawdown is 17.44%, but there is significant variation by year and style. To account for that variation, we style-adjust maximum drawdown within each time period. This adjusted measure, which we abbreviate as MDD, is used throughout our ensuing analyses.

Our first analysis using MDD considers its predictive power with respect to mutual fund performance. We find that an equal weight portfolio of funds in the lowest quintile of past MDD has an alpha ranging from 0.61% (*t*-stat = 1.10) to 1.44% (*t*-stat = 1.94) per year. While suggestive of outperformance in isolation, those alphas are not statistically different from the alphas of an equal weight portfolio of funds in the highest quintile of past MDD. Hence, the initial results suggest that MDD's power to predict fund performance is, at best, debatable.

Nevertheless, the predictive power becomes clear when focusing on mutual funds with relatively strong past performance. While notable prior studies (e.g., Carhart, 1997, and Fama and French, 2010) present evidence that fund performance does not persist, we find that an equal weight portfolio of funds in the lowest quintile of past MDD and highest quintile of past performance has large, positive alphas ranging from 2.09% (*t*-stat = 2.58) to 2.68% (*t*-stat = 2.87) per year.[55] Those alphas are also statistically different from and economically greater than those of an equal weight portfolio of funds in the highest quintile of past MDD and highest quintile of past performance. Given that MDD itself is also persistent, we interpret those positive alphas as evidence of value creation for investors through the utilization of fund managers' risk management skills.

Further performance analysis shows that the extent of value creation through those skills varies considerably over time. Pastor, Stambaugh, and Taylor (2017) demonstrate that the number of profitable opportunities in a given asset market is non-constant. When the number of profitable

[55] Using past performance alone, we do not find evidence of performance persistence in our sample.

opportunities increases, we should expect a mutual fund's performance to improve. Consistent with that logic, when von Reibnitz (2017) cross-sectional return dispersion is high—suggesting a turbulent market with relatively greater opportunities for profit and loss—an equal weight portfolio of funds in the lowest quintile of past MDD and highest quintile of past performance has an alpha ranging from 4.17% (*t*-stat = 1.98) to 5.21% (*t*-stat = 1.71) per year. In all other time periods, the alpha of that same portfolio is statistically indistinguishable from zero. Accordingly, we conclude that risk management skills are, perhaps unsurprisingly, most valuable when markets are turbulent.

Risk management skills could be driven by the ability to select individual stocks or to time the overall market. We determine which aspect MDD is indictive of by regressing the Daniel, Grinblatt, Titman and, and Wermers (1997) "Characteristic Selectivity" (CS) and "Characteristic Timing" (CT) measures on funds' past MDDs. We find a negative relation between MDD and CS but no relation between MDD and CT. In other words, managers with low MDD tend to select stocks better than managers with high MDD but show no difference in market timing. Those results suggest that the risk management skills of low MDD fund managers are driven those managers' ability to select individual stocks rather their ability to time the overall market.

We next estimate the MDD-flow relation to analyze how investors respond to maximum drawdown. In a panel regression setting, we find that, on average, there is a strong, negative relation between MDD and net flows: a 10% increase in MDD is associated with an annualized net outflow of 4.8% (*t*-stat = −6.80). Looking in the cross-section, we would expect the relation to be stronger among a more risk averse clientele. Fittingly, the effect on flows from increases in MDD for funds following an income style—who should attract a relatively risk averse clientele (Blackburn, Goetzmann, and Ukhov, 2009, and Polkovnichenko, Wei, and Zhao, 2019)—is almost three times larger than average. Looking in the time-series, we would also expect the relation to be stronger when the general level of risk aversion in the market is higher. Wang and Young (2020) show that investors are more risk averse when there have been a greater number of recent terrorist attacks, and correspondingly, we find a stronger relation between fund flow and MDD when there has been a greater number of recent terrorist attacks. We further find that the flow-performance relation is flatter for funds with high MDD, which suggests that investors discount outperformance when it is accompanied by high MDD. Given the earlier performance results, this discounting seems rational, as only outperforming funds with low MDD tend to have persistent performance.

Our results contribute to several areas of the mutual fund literature. First, our performance results run counter to “the conventional wisdom that active management does not create value for investors” (Cremers, Fulkerson, and Riley, 2019, pg. 8). We find evidence of risk management skills that create for investors persistently low maximum drawdowns and persistently high alphas. Like other work—such as Amihud and Goyenko’s (2013) selectivity—we find that revealing that value requires accounting for both the measure of interest (in this case, maximum drawdown) and past performance. Our results do not suggest universal value creation for investors through risk management skills, but consistent with Kosowski, Timmermann, Wermers, and White (2006), our results do suggest that a meaningful number of fund managers do use those skills to create value.

Second, we demonstrate that maximum drawdown has a novel effect across multiple mutual fund dimensions. Measures of downside risk have been considered before in the literature (e.g., Artavanis, Eksi, and Kadlec, 2019), as have measures of overall risk (e.g., Jordan and Riley, 2015) and measures of the upside (e.g., Akbas and Genc, 2020). Our results on maximum drawdown are distinctive, however, because they show that (i) the impact of maximum drawdown is not subsumed by those other measures and (ii) the impact of maximum drawdown is present with respect to both performance and flows. The later point makes our findings most like those of

Bodnaruk, Chokaev, and Simonov (2019), who find that fund managers skilled in downside risk timing perform better and receive greater flows.

Third and finally, our flow results speak in favor of the argument that, rather than efficiently process all relevant information before allocating capital, investors tend to focus on the highly salient information. Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016) both show that commonly available CAPM alphas, not multi-factor alphas, best explain net flows, and Evans and Sun (2020) show that investors better accounted for the size and value effects after Morningstar changed their methodology to account for those effects. The large impact of maximum drawdowns on flows fits well in this context, as maximum drawdowns are highly salient. Moreover, unlike the many salient, but seemingly irrational, drivers of flows (e.g., Cooper, Gulen, and Rau, 2005, and Solomon, Soltes, and Sosyura, 2014), the power of maximum drawdown to predict performance suggests that investors are rationally taking advantage of maximum drawdown's wide availability.

3.2. Data and methods

3.2.1. Mutual fund sample

We use the CRSP Survivor-Bias-Free U.S. Mutual Fund database to build our sample of actively managed U.S. equity mutual funds. We drop the funds listed in CRSP as index funds, ETFs, or variable annuities; keep the funds with CRSP objective codes EDCI, EDCS, EDCM, EDYI, EDYB or EDYG; and require that the funds invest at least 70% of their assets in common stocks. We further use Lipper, Strategic Insight, and Wiesenberger investment objective codes to drop the funds not following a long-only strategy. To address incubation bias (Evans, 2010), we exclude a given fund from the sample until it is at least two years old and until it reaches at least $20 million assets. We aggregate the share class level CRSP data into fund level data using the

WFICN variable available in MFLINKS. Fund-level assets are the sum across all share classes. Other fund-level data, such as net return, are asset-weighted averages of the share-level data.

We use daily returns to compute funds' maximum drawdowns. Because daily returns are not available in CRSP until September 1998, we begin our study in 1999. Our final sample, which spans from January 1999 to December 2019, has 2,185 unique funds and 262,921 fund-month observations.

3.2.2. Measuring maximum drawdown

We measure unadjusted maximum drawdown between time 0 and time T as:

$$\text{Unadjusted Maximum Drawdown} = \text{Max}\frac{\mathrm{X}_{t_1} - \mathrm{X}_{t_2}}{\mathrm{X}_{t_1}} \ \text{ s.t. } 0 \le t_1 \le t_2 \le T \qquad (3.1)$$

where X_{t_1} is the cumulative return from time 0 to time t_1. At the beginning of each month, we calculate the unadjusted maximum drawdown of each mutual fund over the past 12 months. This unadjusted value is driven in significant part by a fund's style, rather than a fund's manager's risk management skills. Therefore, to remove the style component, we subtract from each individual unadjusted value the mean unadjusted value during the same time period for all funds with the same style. We use this style-adjusted maximum drawdown (denoted simply as MDD) in all of our subsequent tests.[56]

3.2.3. Risk-adjusting performance

We measure risk-adjusted mutual fund performance using daily fund returns and both the Fama and French (1993) and Carhart (1997) four-factor model (henceforth, the FF4 model) and the Cremers, Petajisto, and Zitzewitz (2012) seven-factor model (henceforth, the CPZ7 model).

[56] This style-adjusted maximum drawdown should be salient to investors because Morningstar reports the MDD of the mutual fund, its style, and its benchmark index.

We thank Ken French for making the FF4 factors available on his website (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The CPZ7 factors are calculated from benchmark returns available in Morningstar. We use net returns throughout our analysis, and to ease interpretation, we annualize the alphas from these models in our tables.

3.3. The maximum drawdowns of mutual funds

We begin our analyses by considering the size of the maximum drawdowns within our sample. Figure 3.1 plots the average unadjusted maximum drawdown in each year for the full sample of mutual funds and for subsamples of growth funds and income funds. As expected, there are significant fluctuations in unadjusted maximum drawdowns from 1999 to 2018. Full sample averages range from 50.4% in 2008 to 4.50% in 2017. There are also significant fluctuations between styles. The average unadjusted maximum drawdown of growth funds in 2001 was 12.85% greater than that of income funds, while in 2009 that difference was −3.41%. This variation shows the importance of our subsequent use of style-adjusted maximum drawdown.

Table 3.1, Panel A provides the characteristics of mutual funds in the highest and lowest quintiles of style-adjusted maximum drawdown. Low MDD funds, on average, have higher returns, higher alphas, and lower volatilities. They also tend to have lower turnover and expense ratios. If, as in Panel B, we focus on just the funds that are also in the top quintile of past performance—measured over the previous 24 months using the CPZ7 model—low MDD funds continue to have lower average volatilities, turnover ratios, and expense ratios; however, there are no statistically significant differences in returns or alphas.

3.4. Can maximum drawdown predict mutual fund performance?

3.4.1. Returns and alphas

We next consider the relation between past MDD and subsequent performance. Motivated by Cremer's (2017) model of successful active management and empirical results from prior work (e.g., Amihud and Goyenko, 2013), we place a particular emphasis on the MDDs of mutual funds with relatively strong past performance. For these tests, we create a set of portfolios based on past performance and past MDD. We form those portfolios by first unconditionally sorting funds into quintiles at the beginning of each month based on their past 24-month CPZ7 alphas and past 12-month MDDs. We then equal weight the funds in the resulting twenty-five groups.

Our analysis of these portfolios begins by studying raw returns. Figure 3.2 shows the cumulative dollar returns on the five portfolios formed using the highest quintile of past performance. As past MDD decreases, the final portfolio values monotonically increase. A one-dollar investment made at the beginning of 2000 in the portfolio of the mutual funds in the highest past performance quintile and lowest past MDD quintile is, at the end of 2019, worth \$6.80. In comparison, the portfolio of the funds in the highest past performance quintile and highest past MDD quintile is, at the end of 2019, worth only \$2.43.

This first test suggests MDD has predictive power with respect to performance, but it is far from conclusive. Raw returns do not account for the potentially large differences in factor exposure between the portfolios. To evaluate the risk-adjusted performance of the portfolios, we turn to the FF4 and CPZ7 models. Table 3.2 shows the alphas for all of the portfolios using each model (FF4 in Panel A and CPZ7 in Panel B). There is some evidence that, unconditionally, mutual funds in the lowest past MDD quintile outperform—the FF4 alpha for the portfolio of those funds is 1.44% per year (t-stat = 1.94)—but, using either model, the difference in alpha between the highest and lowest past MDD quintile portfolios is not statistically different from zero. The evidence is much stronger after conditioning on past performance. The portfolio of the funds in the highest past performance quintile and lowest past MDD quintile has an FF4 alpha of 2.68% per year (*t*-stat = 2.87) and a CPZ7 alpha of 2.09% per year (*t*-stat = 2.58). Moreover, that portfolio significantly outperforms (i) the portfolio formed using the funds from the top past performance quintile, but highest past MDD quintile, and (ii) the portfolio formed using the funds from the lowest past MDD quintile, but lowest past performance quintile.

3.4.2. Subsequent MDD

If MDD is related to risk management skills, we would expect a fund's MDD to persist. Table 3.3 examines the out-of-sample MDDs of the prior sub-section's portfolios. From the results, it is clear that MDD persists. Setting aside past performance, the portfolio of the mutual funds in the lowest past MDD quintile has a subsequent average MDD of 14.53%, compared to 20.28% for the portfolios of the funds in the highest past MDD quintile. Within each quintile of past performance, there is a similar difference in out-of-sample MDD between the lowest and highest past MDD quintiles. The consistency of these results suggests that a low MDD is a not a random event, but rather an indicator of risk management skills.

3.4.3. Market conditions

The previous results suggest that the mutual funds in the highest past performance quintile and lowest past MDD quintile have large subsequent alphas and low subsequent MDDs, which is an appealing combination. The performance of those funds, however, could vary substantially over time. As modeled by Pastor, Stambaugh, and Taylor (2017), fund managers are constrained by the non-constant number of profitable opportunities in their asset market. When the number of profitable opportunities increases, we should expect a given fund's performance to increase. We proxy for the number of profitable opportunities (or, alternatively, the turbulence of the market) using the von Reibnitz (2017) measure of cross-sectional return dispersion (RD). Von Reibnitz (2017) calculates RD in a given month as the standard deviation of the monthly returns on individual S&P 500 stocks. An increase in RD implies an increase in the number of profitable opportunities.

In Table 3.4, we first use one-month lagged RD to sort our time period in quintiles. Then, within each of the time period quintiles, we estimate the performance of the portfolio of the mutual funds in the highest past performance quintile and lowest past MDD quintile. As shown, the alpha of that portfolio varies substantially depending on market conditions, with alpha decreasing near monotonically as RD decreases. Within the highest RD quintile, the FF4 alpha is 5.21% per year (*t*-stat = 1.71), and the CPZ7 alpha is 4.17% per year (*t*-stat = 1.98). There are, conversely, no statistically significant alphas in the other market conditions. Within the lowest RD quintile, the FF4 alpha is −0.88% per year (*t*-stat = −1.14), and the CPZ7 alpha is −1.21% per year (*t*-stat = −1.20). In untabulated analysis, we find that this portfolio has a relatively low out-of-sample MDD regardless of market conditions, but with respect to overall risk-adjusted performance, these results indicate significant state dependence.

3.4.4. What aspect of skill is MDD capturing?

Next, we investigate the relation between MDD and stock picking and market timing abilities. Either ability could be the channel through which risk management skills operate. A fund manager could be skilled at selecting the correct stocks to produce a portfolio with a low expected MDD, or they could be skilled at knowing when to limit the portfolio's overall market exposure. Potentially, a fund manager could use both skills simultaneously.

Daniel, Grinblatt, Titman and, and Wermers (1997) use mutual fund holdings data to generate two measures—"Characteristic Selectivity" (CS) and "Characteristic Timing" (CT)—that quantify the fund performance that comes from stock picking and market timing. If the managers of funds with low MDD have stock picking ability, then MDD should have a negative relation with CS. Likewise, if the managers of funds with low MDD have market timing ability, then MDD should have a negative relation with CT.

We estimate the relation between MDD and CS using the model described in Eq. (3.2):

$$\mathrm{CS}_{i,t} = \beta_0 + \beta_1 \mathrm{MDD}_{i,t-1} + \beta_2 \mathrm{Perf}_{i,t-1} + \sum_{j=1}^{J} \gamma_j \mathrm{c}_{j,i,t-1} + \mathrm{FE} + \varepsilon_{i,t} \tag{3.2}$$

$\mathrm{CS}_{i,t}$ is the annualized CS measure of fund *i* in month *t* and is computed following Daniel, Grinblatt, Titman and, and Wermers (1997).[57] $\mathrm{MDD}_{i,t-1}$ is the style-adjusted percentage maximum drawdown of fund i over the twelve months ending with month $t-1$. $\mathrm{Perf}_{i,t-1}$ is the performance of fund i over the twelve months ending with month $t-1$. We measure performance as the percentile ranking of the within-style cumulative net return, with a 100 given to the best performing fund and a zero given to the worst. $\mathrm{c}_{j,i,t-1}$ is the jth characteristic of fund i as of the end of month $t-1$. Characteristics in the model include the natural logarithm of fund age, the natural logarithm of fund assets, the natural logarithm of fund family assets, the expense ratio, and the turnover ratio. FE represents style fixed effects, and year-month fixed effects. The relation between MDD and CT is estimated analogously.

Table 3.5 shows results from estimates of that model. Column (1) shows a negative relation between MDD and CS. A 10% increase in MDD is associated with a 1% decrease in annualized CS (*t*-stat = −2.31). That result suggests that the managers of funds with low MDD have greater

[57] We use the Thomson Reuters database for mutual fund holdings and the CRSP database for information on stocks.

stock selection skill than managers of funds with high MDD. Conversely, column (2) shows no significant relation between MDD and CT. The coefficient associated with MDD is economically and statistically zero, implying there is no difference in market timing skill between managers of funds with low and high MDD. Based on these results, we therefore conclude that the risk management skills of low MDD fund managers are primarily related to their ability to select individual stocks, not their ability to time the overall market.

3.4.5. Uniqueness of MDD

Finally, we consider the utility of MDD as a performance predictor relative to other related measures. Our goal is to determine whether the predictive power of MDD is unique or if its power is subsumed by other previously analyzed measures. We focus on three previously analyzed measures: Jordan and Riley (2015) volatility (Vol), Artavanis, Eksi, and Kadlec (2019) downside beta (DBeta), and Akbas and Genc (2020) maximum return (MAX). With respect to performance, Vol is of the most interest because it alone among these measures has been previously shown to predict performance. The other measures are of more interest with respect to flows.

To establish a baseline, Panel A of Table 3.6 shows the alphas of equal weight portfolios formed using the mutual funds in the highest past performance quintile and a specific quintile of each of the measures. We use the lowest quintiles for MDD, Vol, and DBeta, and the highest quintile for MAX. Consistent with Jordan and Riley (2015), the portfolio of the funds in the lowest past Vol quintile does outperform, and consistent with Akbas and Genc (2020), the portfolio of the funds in the highest past MAX quintile does not outperform. Although Artavanis, Eksi, and Kadlec (2019) only focus on fund flows, we find DBeta also predicts performance. The portfolio of the funds in the lowest past DBeta quintile has an FF4 alpha of 2.46% per year (*t*-stat = 2.59) and a CPZ7 alpha of 1.32% per year (*t*-stat = 1.88).

Based on those results, it is reasonable to question whether MDD still has predictive power after accounting for Vol or DBeta. Table 3.6, Panel B speaks to that question by evaluating the performance of mutually exclusive portfolios. To explain by example, the 'MDD excl. Vol' portfolio consists only of the mutual funds in Panel A's MDD-based portfolio that are not in Panel A's Vol-based portfolio, and the 'Vol excl. MDD' portfolio consists only of the funds in the Vol-based portfolio that are not in MDD-based portfolio. Outperformance from the 'MDD excl. Vol' portfolio indicates that MDD has unique predictive power relative to Vol because the predictive power from Vol should be absent from that portfolio.

The mutually exclusive portfolios show that the predictive power of MDD is unique. Each of the 'MDD excl.' portfolios has an economically large and statistically significant alpha. For example, the 'MDD excl. Vol' portfolio has an FF4 alpha of 2.38% per year (*t*-stat = 2.18) and a CPZ7 alpha of 2.08% per year (*t*-stat = 2.08). MDD does not subsume the predictive power of Vol or DBeta, but these results suggest that MDD is able to detect outperforming mutual funds that are undetected by previously analyzed measures.

3.5. Do investors respond to maximum drawdown?

3.5.1. MDD-flow relation

Investors are generally risk averse. Consequently, it is reasonable to expect that, all other things equal, investors would prefer mutual funds that have not experienced large maximum drawdowns. If our previous performance results had shown that large maximum drawdowns in the past were predictive of outperformance in the future, that expectation would have to re-evaluated; however, we found that low maximum drawdowns in the past are actually predictive of outperformance in the future—which further cements the initial expectation. We investigate how investors respond to maximum drawdowns in reality by testing the relation between net fund flows and MDD. Our general model is:

$$\mathrm{Flow}_{i,t} = \lambda_0 + \lambda_1 \mathrm{MDD}_{i,t-1} + \lambda_2 \mathrm{Perf}_{i,t-1} + \sum_{j=1}^{J} \delta_j \mathrm{c}_{j,i,t-1} + \mathrm{FE} + \varepsilon_{i,t} \qquad (3.3)$$

where $\mathrm{Flow}_{i,t}$ is the annualized percentage net flow of fund i in month t. We calculate the net flow following Sirri and Tufano (1998). $\mathrm{MDD}_{i,t-1}$ is the style-adjusted percentage maximum drawdown of fund i over the twelve months ending with month $t-1$. $\mathrm{Perf}_{i,t-1}$ is the performance of fund i over the twelve months ending with month t – 1. We measure performance as the within-style cumulative net return percentile ranking, with a 100 given to the best performing fund and a zero given to the worst. To control for potential non-linearity in the flow-performance relation, we convert that measure of performance into piecewise performance measures styled after Sirri and Tufano (1998) in some model specifications. $\mathrm{c}_{j,i,t-1}$ is the jth characteristic of fund i as of the end of month t – 1. Characteristics in the model include the natural logarithm of fund age, the natural logarithm of fund assets, the natural logarithm of fund family assets, the expense ratio, and the turnover ratio. FE represents style fixed effects, and year-month fixed effects.

Table 3.7 presents results from estimates of this model. Columns (1) and (2) show that, consistent with our initial expectation, flows have a negative relation with MDD. A 10% increase in MDD is associated with an annualized outflow of 4.8% (t-stat = −6.80) to 5.2% (t-stat = −7.38). In columns (3) and (4), we consider the expectation that the relation will be stronger among more risk averse investors. Relative to other equity styles, income funds should attract a more risk averse clientale; hence, we expect income funds to have a stronger flow response to MDD. We test that hypothesis by interacting MDD with a dummy variable (Income) that is equal to one if the fund's investment style is income. As shown, the response is stronger for income funds: a 10% increase in MDD for income funds is associated with an annualized outflow 8.4% (t-stat = −3.06) to 8.7% (t-stat = −3.04) greater than that of funds following other styles. Therefore, we conclude that the relation between flows and MDD is stronger among more risk averse investors.

While investors vary cross-sectionally in their risk aversion, the general level of investor risk aversion also varies over time. Wang and Young (2020) show that the general level of investor risk aversion is higher when there have been a greater number of recent terrorist attacks, so we expect that the strength of the negative flow-MDD relation will vary with that trend. Specifically, when there have been a greater number of recent terrorist attacks, the flow-MDD relation will be stronger. To test that hypothesis, we interact MDD with the z-score of the number of recent terrorist attacks in the world that involve human casualties or injuries.[58] Column (5) shows that flow-MDD relation is indeed stronger when there have been a greater number of recent terrorist attacks. A one standard deviation increase in attacks increases the coefficient associated with MDD by about 20% (=−0.11/−0.56) to 23% (=−0.12/−0.52). From these results, we conclude that the strength of the flow-MDD relation does vary depending on the general level of investor risk aversion.

The final two columns of Table 3.7 consider how MDD effects the shape of the flow-performance relation. We expect that investors will discount highly ranked performance if that performance is accompanied by large drawdown risk. We test this hypothesis by interacting our performance measures with a dummy variable (High MDD) that is equal to one if MDD is within the highest MDD quintile. As shown in column (8), the flow-performance relation is flatter for mutual funds with high MDD. Compared to other funds, the flows of funds with high MDD respond relatively more to lowly ranked performance (Low Perf) and relatively less to highly ranked performance (High Perf). Among funds with highly ranked performance, when MDD is

[58] Data on terrorist attacks is downloaded from the University of Maryland's Global Terrorism Database (DTD).

high the responsiveness of flows to an increase in performance is reduced by 26% (=0.26/0.99). Thus, consistent with our expectation, investors do discount the highly ranked performance of funds subject to large drawdown risk.

3.5.2. MDD uniqueness

We finally consider how the MDD-flow relation is affected by the inclusion of our earlier set of previously analyzed measures (Vol, DBeta, and MAX). Our general model is the same as in the previous sub-section, except for the inclusion of those variables. The goal of this test is to determine whether the impact of MDD on flows is unique or if it is subsumed by already known relations.

Table 3.8 shows the results from this test. Columns (1) through (3) show that the impact of MDD on flows is not subsumed by Vol, but that the impact of Vol is subsumed by MDD. Jordan and Riley (2015) did not consider the Vol-flow relation, so these results do not overturn previous results, but they are nevertheless notable. Given that Vol and MDD have similar saliency—they are listed on the same page on Morningstar's website—this result suggests investors weight the specific component of risk captured by MDD more heavily than the more general component of risk captured by Vol. Conversely, columns (4) through (9) show that MDD's impact on flows is not subsumed by DBeta or MAX, and that DBeta's and MAX's impacts are not subsumed by MDD. These results suggest that investors, as a group, are not myopic when allocating capital. They consider multiple components of each fund's history.

3.6. Conclusion

We evaluate the importance of mutual fund managers' risk management skills—as quantified by styled-adjusted maximum drawdown. We find that a fund's past maximum drawdown has unique predictive power with respect to subsequent performance and that investors give considerable weight to a fund's past maximum drawdown when allocating capital. Among funds with relatively strong past performance, those with relatively low past maximum drawdowns have an average FF4 alpha of 2.68% per year (*t*-stat = 2.87). That result, in conjunction with the persistence in funds' maximum drawdowns, suggests that having a low maximum drawdown is a signal of fund manager risk management skills. The gains to investors from those skills tend to vary over time, with most of the gains being realized during turbulent market conditions. Further, we find that low MDD is associated with stock picking skill but not market timing skill.

Given those results, it is not surprising that we find that (i) there is a strong negative relation between fund flows and maximum drawdowns, particularly among more risk averse investors and during periods of generally heightened risk aversion, and (ii) investors discount the outperformance of funds that have had relatively large maximum drawdowns. Considered as a whole, our paper demonstrates that mutual fund managers' risk management skill—an important, but understudied, skill in the literature—has a significant and unique influence on mutual fund performance and flows.

Table 3.1: Characteristics of mutual funds with low and high MDD

This table presents mean characteristics for mutual funds in the top and bottom quintiles of past style-adjusted maximum drawdown (MDD). Panel A shows the characteristics for all funds, unconditional on past performance. Panel B shows the characteristics after excluding all funds not in the top quintile of past performance. The *t*-statistics associated with the differences are calculated using standard errors clustered on fund and year-month. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

Panel A: Full Sample

	Low MDD	High MDD	Difference	*t*-statistic	All
Unadjusted MDD (%)	13.34	23.27	-9.93***	-19.48	17.44
MDD (%)	-4.59	5.51	-10.10***	-20.78	0.00
Return (%)	12.26	4.37	7.89***	9.15	8.34
Volatility (%)	16.96	22.42	-5.46***	-12.70	19.08
Age (months)	212.44	205.04	7.41	1.25	218.21
Asset (millions)	1,780.23	1,155.29	624.93***	2.97	1,669.45
Family Asset (millions)	90,335.90	82,591.49	7,744.40	0.80	103,913.10
Expense (%)	1.16	1.30	-0.14***	-9.54	1.18
Turnover (%)	61.77	94.39	-32.62***	-11.89	76.05
FF4 Alpha (%)	1.08	-2.47	3.55***	9.96	-0.79
CPZ7 Alpha (%)	1.14	-2.46	3.60***	12.29	-0.85

Panel B: Top Performing Funds

	Low MDD	High MDD	Difference	*t*-statistic	All
Unadjusted MDD (%)	13.97	22.9	-8.93***	-10.73	16.95
MDD (%)	-5.03	5.79	-10.82***	-18.28	-1.00
Return (%)	14.69	14.49	0.20	0.11	14.28
Volatility (%)	17.96	23.59	-5.63***	-6.46	19.75
Age (months)	206.49	188.63	17.87**	2.18	205.58
Asset (millions)	1,900.06	1,420.05	480.01*	1.81	1,871.08
Family Asset (millions)	88,178.14	77,335.76	10,842.38	0.79	103,608.30
Expense (%)	1.18	1.33	-0.16***	-6.27	1.21
Turnover (%)	56.87	91.84	-34.96***	-8.82	70.13
FF4 Alpha (%)	5.26	4.73	0.53	0.63	4.67
CPZ7 Alpha (%)	6.13	6.70	-0.57	-0.91	5.91

Table 3.2: Alphas of portfolios formed based on past performance and MDD

This table shows the alphas of equal weight portfolios formed by the unconditional quintile sorting of mutual funds based on past performance and past style-adjusted maximum drawdown (MDD). Panel A measures the portfolios' alphas using the FF4 model, while Panel B uses the CPZ7 model. The alphas are measured over the time period January 2000 to December 2019. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

Panel A: FF4 Alpha

	MDD Quintile						
Alpha Quintile	1 (Low)	2	3	4	5 (High)	Low − High	All
5 (High)	2.68***	1.17	0.55	0.12	-2.25*	4.93***	0.77
	[2.87]	[1.54]	[0.71]	[0.12]	[-1.67]	[3.04]	[1.16]
4	1.13	1.87	-0.38	-1.23*	-2.11*	3.24**	0.22
	[1.64]	[1.32]	[-0.80]	[-1.68]	[-1.65]	[2.21]	[0.38]
3	1.85	-0.35	-0.48	-1.08*	-2.57**	4.42**	-0.52
	[1.14]	[-0.70]	[-0.98]	[-1.68]	[-2.32]	[2.20]	[-1.08]
2	-0.08	-0.70	-1.36**	-1.80***	-2.62**	2.54**	-1.36**
	[-0.10]	[-1.16]	[-2.54]	[-3.27]	[-2.46]	[2.00]	[-2.58]
1 (Low)	-0.13	-1.55**	-1.68**	-1.83**	0.92	-1.05	-0.09
	[-0.17]	[-2.15]	[-2.27]	[-2.49]	[0.21]	[-0.23]	[-0.04]
High − Low	2.81***	2.72***	2.24***	1.95**	-3.17		0.85
	[3.15]	[3.72]	[2.76]	[2.02]	[-0.72]		[0.38]
All	1.44*	0.28	-0.69	-1.26**	-0.75	2.19	-0.20
	[1.94]	[0.45]	[-1.37]	[-2.18]	[-0.31]	[0.86]	[-0.29]

Panel B: CPZ7 Alpha

	MDD Quintile						
Alpha Quintile	1 (Low)	2	3	4	5 (High)	Low − High	All
5 (High)	2.09**	0.80	0.59	-0.03	-1.46	3.55**	0.70
	[2.58]	[1.41]	[1.01]	[-0.04]	[-1.21]	[2.16]	[1.51]
4	0.37	1.04	-0.66*	-1.59***	-2.21**	2.58*	-0.30
	[0.73]	[0.97]	[-1.84]	[-2.67]	[-2.03]	[1.88]	[-0.81]
3	0.42	-0.83**	-0.93**	-1.49***	-2.51**	2.93*	-1.05***
	[0.39]	[-2.02]	[-2.56]	[-2.97]	[-2.57]	[1.75]	[-3.50]
2	-1.10*	-1.33***	-1.82***	-2.04***	-2.44***	1.35	-1.79***
	[-1.84]	[-2.84]	[-4.28]	[-4.57]	[-2.68]	[1.11]	[-4.90]
1 (Low)	-0.92	-1.89***	-2.22***	-2.14***	1.39	-2.31	-0.23
	[-1.22]	[-3.05]	[-3.80]	[-3.76]	[0.30]	[-0.48]	[-0.10]
High − Low	3.00***	2.69***	2.80***	2.11**	-2.85		0.93
	[3.31]	[3.75]	[3.39]	[2.46]	[-0.61]		[0.39]
All	0.61	-0.27	-1.06***	-1.56***	-0.39	0.99	-0.53
	[1.10]	[-0.63]	[-3.15]	[-3.87]	[-0.16]	[0.38]	[-1.01]

Table 3.3: MDDs of portfolios formed based on past performance and MDD

This table shows the average annual unadjusted maximum drawdowns of equal weight portfolios formed by the unconditional quintile sorting of mutual funds based on past performance and past MDD. The portfolio MDDs are measured over the time period 2000 to 2019. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

	MDD Quintile						
Alpha Quintile	1 (Low)	2	3	4	5 (High)	Low − High	All
5 (High)	14.80***	15.98***	17.25***	18.47***	21.32***	-6.52***	16.83***
	[6.70]	[6.48]	[6.77]	[6.60]	[6.83]	[-4.16]	[6.93]
4	14.45***	15.36***	16.24***	17.44***	20.11***	-5.66***	15.98***
	[6.46]	[6.50]	[6.55]	[6.51]	[6.61]	[-3.86]	[6.62]
3	14.45***	15.34***	16.02***	16.89***	19.83***	-5.38***	15.98***
	[6.56]	[6.36]	[6.35]	[6.41]	[6.55]	[-3.72]	[6.49]
2	14.39***	15.31***	16.04***	17.10***	19.93***	-5.54***	16.50***
	[6.40]	[6.40]	[6.44]	[6.48]	[6.73]	[-4.11]	[6.49]
1 (Low)	14.55***	16.00***	16.68***	17.39***	20.37***	-5.82***	17.74***
	[6.41]	[6.40]	[6.45]	[6.48]	[6.86]	[-3.92]	[6.59]
High − Low	0.25	-0.03	0.57	1.07**	0.95		-0.91
	[0.59]	[-0.07]	[1.00]	[2.32]	[1.70]		[-1.72]
All	14.53***	15.49***	16.24***	17.27***	20.28***	-5.76***	16.54***
	[6.59]	[6.44]	[6.48]	[6.48]	[6.73]	[-4.05]	[6.62]

Table 3.4: Performance as a function of market conditions

This table shows the FF4 and CPZ7 alphas of a portfolio formed by the unconditional quintile sorting of mutual funds based on past performance and past style-adjusted maximum drawdown (MDD). The portfolio evaluated consists of the funds in both the lowest past MDD quintile and the highest past performance quintile. The time period of January 2000 to December 2019 is divided into quintiles based on one-month lagged von Reibnitz (2017) cross-sectional return dispersion (RD). *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

RD Quintile	FF4 Alpha	CPZ7 Alpha
5 (High)	5.21*	4.17*
	[1.71]	[1.98]
4	2.50	2.92
	[1.13]	[0.95]
3	0.21	0.28
	[0.24]	[0.31]
2	0.14	0.58
	[0.13]	[0.61]
1 (Low)	-0.88	-1.21
	[-1.14]	[-1.20]
High − Low	6.09*	5.38**
	[1.93]	[2.31]
All	2.68***	2.09**
	[2.87]	[2.58]

Table 3.5: What aspects of skill is MDD capturing?

This table shows estimates of Eq. (3.2), which is our model of the relation between "Characteristic Selectivity" (CS) and "Characteristic Timing (CT)" measures of Daniel, Grinblatt, Titman and, and Wermers (1997) and lagged style-adjusted maximum drawdown (MDD). All the results include lagged fund characteristics as controls and use style fixed effects, and year-month fixed effects. The standard errors are clustered by fund and year-month. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

	(1) CS	(2) CT
MDD	-0.10**	-0.01
	[-2.31]	[-0.42]
Ln(Age)	0.01	0.01
	[0.29]	[0.36]
Ln(Assets)	-0.07***	0.00
	[-2.61]	[0.02]
Ln(Family Assets)	0.04**	0.00
	[1.98]	[0.42]
Expense	-0.01	0.02
	[-0.06]	[0.30]
Turnover	-0.00	-0.00
	[-0.87]	[-0.35]
Perf	0.00	0.00
	[0.79]	[1.11]
Fixed Effects	Yes	Yes
Observations	219,062	219,062
R-squared	0.05	0.11

Table 3.6: Comparison of MDD and other measures as performance predictors

This table shows, in Panel A, the FF4 and CPZ7 alphas of equal weight portfolios formed by the unconditional quintile sorting of mutual funds based on past performance and another measure. The ‘Low MDD’, ‘Low Vol’, and ‘Low DBeta’ portfolios use the funds in both the lowest quintiles of those respective measures and the highest quintile of past performance. The ‘High MAX’ portfolio uses the funds in both the highest quintile of MAX and the highest quintile of past performance. Panel B shows the alphas for mutually exclusive equal weight portfolios formed based on the overlap between the funds in the preceding portfolios. ‘MDD excl. Vol’, for example, consists of the funds in the ‘Low MDD’ portfolio that are not in the ‘Low Vol’ portfolio. The ‘Distinctive’ row reports the percentage of funds that are included in only that particular portfolio among the set of funds that are included in at least one of the two portfolios. All of the alphas in both panels are measured over the time period January 2000 to December 2019. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

Panel A: Performance of Individual Portfolios

	Low MDD	Low Vol	Low DBeta	High MAX
FF4 Alpha	2.68***	2.83***	2.46**	0.42
	[2.87]	[3.15]	[2.59]	[0.48]
CPZ7 Alpha	2.09**	1.71**	1.32*	0.78
	[2.58]	[2.53]	[1.88]	[1.17]

Panel B: Performance of Mutually Exclusive Portfolios

	MDD excl. Vol	Vol excl. MDD	MDD excl. DBeta	DBeta excl. MDD	MDD excl. MAX	MAX excl. MDD
FF4 Alpha	2.38**	3.19***	2.97**	1.81*	2.50**	-0.74
	[2.18]	[3.06]	[2.22]	[1.80]	[2.41]	[-0.66]
CPZ7 Alpha	2.08**	1.94***	2.68**	0.77	1.90**	-0.05
	[2.08]	[2.62]	[2.15]	[1.03]	[2.08]	[-0.06]
Distinctive	48.12%	18.38%	41.87%	25.15%	36.47%	41.42%

Table 3.7: MDD-flow relation

This table shows estimates of Eq. (3.3), which is our model of the relation between net mutual fund flows and style-adjusted maximum drawdown (MDD). For brevity, the variables related to fund characteristics are suppressed in the table. Columns (5) and (6) use style fixed effects, and year fixed effects. The other columns use style fixed effects, and year-month fixed effects. The standard errors are clustered by fund and year-month. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MDD	-0.48***	-0.52***	-0.45***	-0.49***	-0.52***	-0.56***		
	[-6.80]	[-7.38]	[-6.38]	[-6.97]	[-7.37]	[-7.77]		
MDD * Income			-0.87***	-0.84***				
			[-3.04]	[-3.06]				
MDD * Attack					-0.12*	-0.11		
					[-1.84]	[-1.57]		
Attack					0.77	0.76		
					[0.69]	[0.68]		
Perf	0.29***		0.29***		0.29***		0.31***	
	[30.15]		[30.01]		[29.92]		[32.02]	
Low Perf		0.43***		0.42***		0.41***		0.33***
		[12.25]		[12.17]		[11.98]		[8.11]
Mid Perf		0.19***		0.19***		0.19***		0.21***
		[19.10]		[19.07]		[19.22]		[21.32]
High Perf		0.93***		0.93***		0.93***		0.99***
		[19.81]		[19.80]		[19.87]		[20.05]
Perf * High MDD							-0.02*	
							[-1.75]	
Low Perf * High MDD								0.24***
								[3.88]
Mid Perf * High MDD								-0.04***
								[-2.58]
High Perf * High MDD								-0.26***
								[-2.65]
High MDD							-2.19***	-6.27***
							[-3.63]	[-5.96]
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	238,292	238,292	238,292	238,292	238,292	238,292	238,292	238,292
R-squared	0.11	0.12	0.11	0.12	0.10	0.10	0.11	0.12

Table 3.8: Relation between flows, MDD, and MDD-related measures

This table shows estimates from an expansion of our Eq. (3.3), which is our model of the relation between net mutual fund flows and style-adjusted maximum drawdown (MDD). In this instance, we also add Vol, DBeta, and Max to the model. For brevity, the variables related to fund characteristics are suppressed in the table. The standard errors are clustered by fund and year-month. *t*-statistics are in brackets below each coefficient. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
MDD		-0.42***	-0.43***		-0.16*	-0.20**		-0.55***	-0.55***
		[-4.29]	[-4.34]		[-1.92]	[-2.38]		[-7.39]	[-7.59]
Vol	-0.42***	-0.09	-0.14						
	[-4.19]	[-0.63]	[-0.95]						
DBeta				-12.03***	-11.04***	-10.97***			
				[-6.57]	[-5.46]	[-5.44]			
MAX							0.47***	0.75***	0.37***
							[3.29]	[5.19]	[2.58]
Perf	0.31***	0.29***		0.30***	0.29***		0.30***	0.27***	
	[33.70]	[31.63]		[32.70]	[30.95]		[34.03]	[28.91]	
Low Perf			0.44***			0.46***			0.43***
			[12.93]			[13.15]			[12.42]
Mid Perf			0.19***			0.19***			0.18***
			[20.40]			[19.40]			[18.86]
High Perf			0.94***			0.93***			0.90***
			[19.62]			[20.04]			[19.09]
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	238,292	238,292	238,292	238,292	238,292	238,292	238,292	238,292	238,292
R-squared	0.11	0.11	0.12	0.12	0.12	0.12	0.11	0.11	0.12

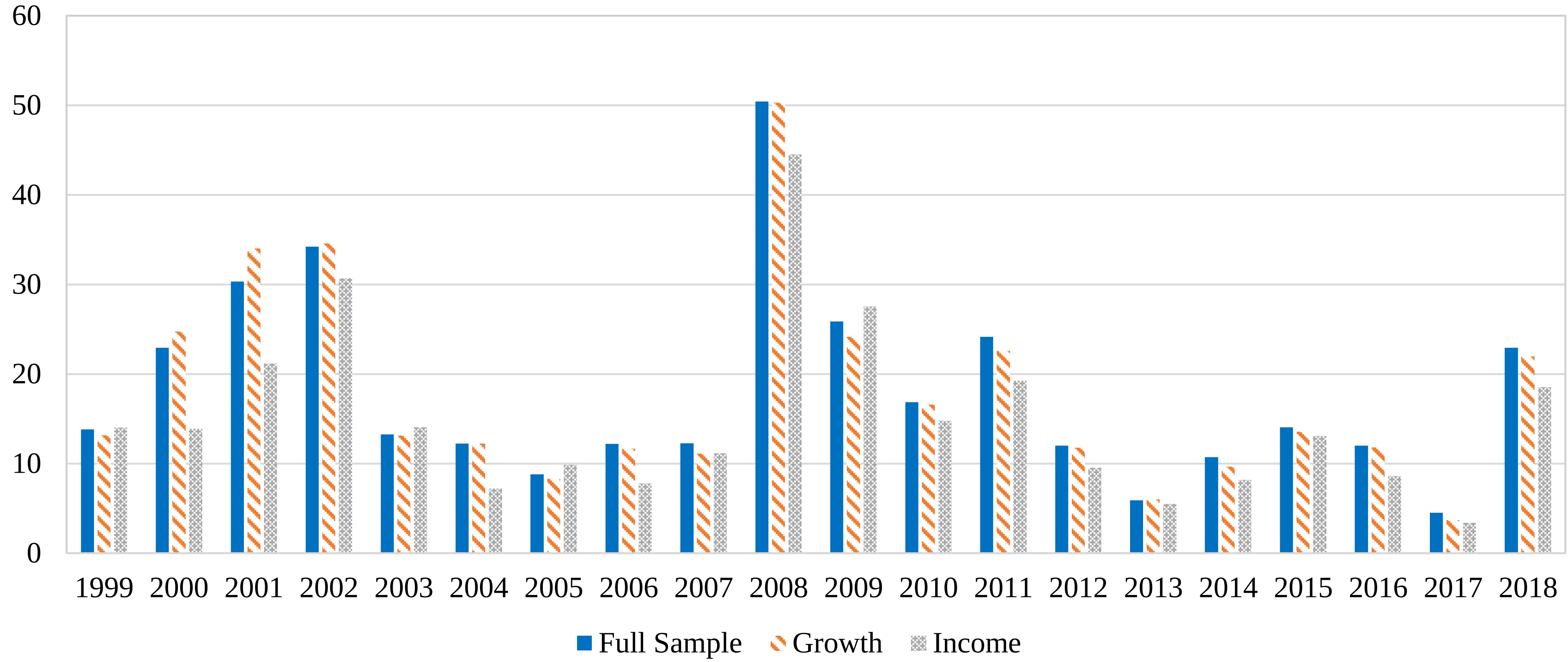

Figure 3.1: Average unadjusted maximum drawdown of mutual funds
This figure presents the average unadjusted maximum drawdown for all sample mutual funds. It also presents results for subsamples of growth funds and income funds. Separate averages are calculated for each year from 1999 to 2018.

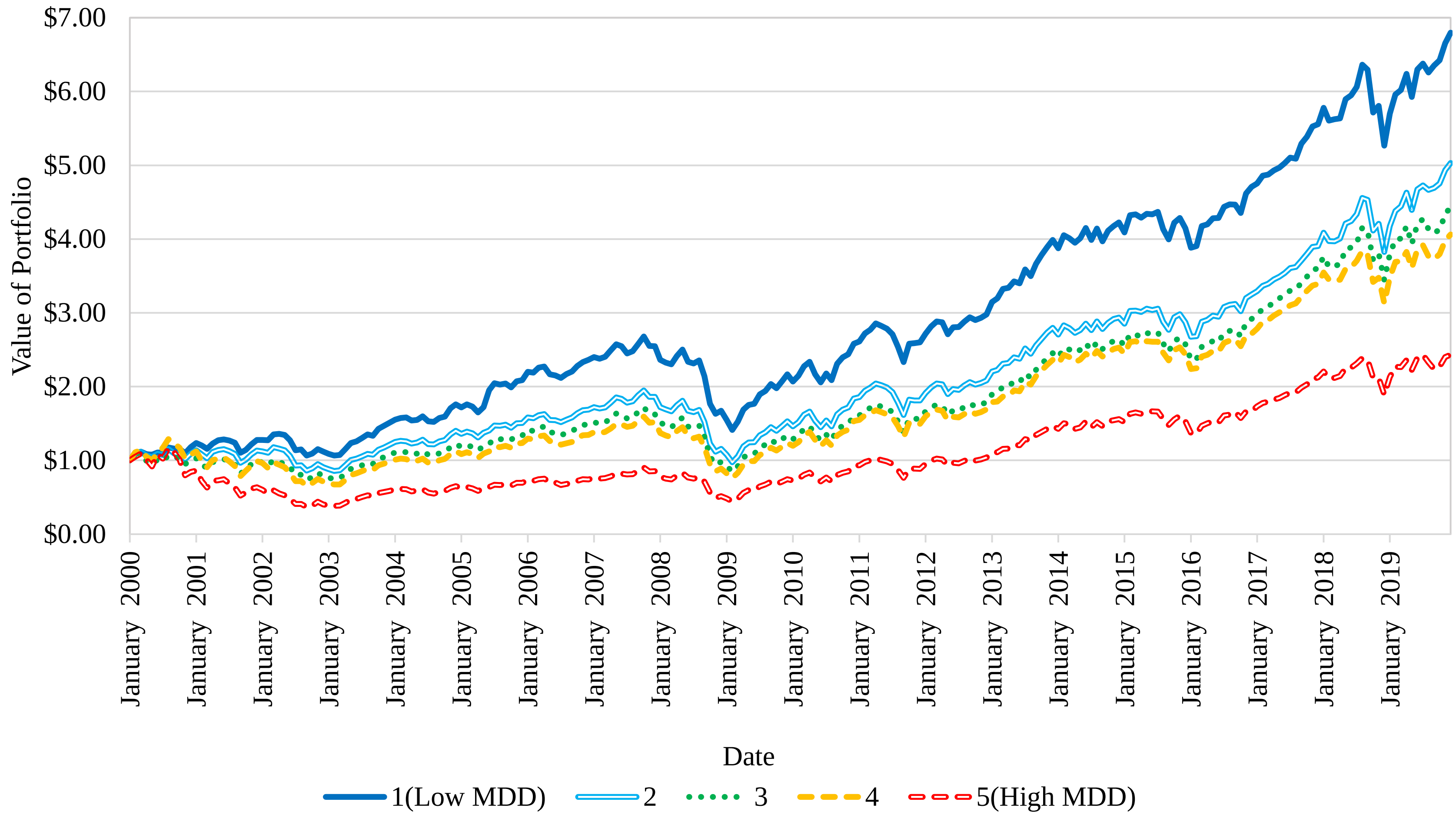

Figure 3.2: Cumulative return on portfolios formed based on past performance and MDD

This figure shows the changing value of a $1 investment made in five different portfolios from January 2000 to December 2019. The portfolios are equal weighted and formed by sorting mutual funds unconditionally on past performance and past style-adjusted maximum drawdown (MDD). The 'Low MDD' portfolio consists of the funds in the highest quintile of past performance and lowest quintile of past MDD. The 'High MDD' portfolio consists of the funds in the highest quintile of past performance and highest quintile of past MDD.

References

Adams, John, Darren Hayunga, and Sattar Mansi. "Diseconomies of scale in the actively-managed mutual fund industry: What do the outliers in the data tell us?" *Critical Finance Review* 7, no.2 (2018), 273-329.

Agarwal, Vikas, and Narayan Y. Naik. "Multi-period performance persistence analysis of hedge funds." *Journal of Financial and Quantitative Analysis* 35, no. 3 (2000): 327-342.

Agarwal, Vikas, and Narayan Y. Naik. "Risks and portfolio decisions involving hedge funds." *The Review of Financial Studies* 17, no. 1 (2004): 63-98.

Agarwal, Vikas, Clifton Green, and Honglin Ren. "Alpha or beta in the eye of the beholder: What drives hedge fund flows?." *Journal of Financial Economics* 127, no. 3 (2018): 417-434.

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik. "Role of managerial incentives and discretion in hedge fund performance." *The Journal of Finance* 64, no. 5 (2009): 2221-2256.

Agarwal, Vikas, Nicole Boyson, and Narayan Naik. "Hedge funds for retail investors? An examination of hedged mutual funds." *Journal of Financial and Quantitative Analysis* 44, no. 2 (2009): 273-305.

Aiken, Adam, Christopher Clifford, and Jesse Ellis. "Out of the dark: Hedge fund reporting biases and commercial databases." *The Review of Financial Studies* 26, no. 1 (2013): 208-243.

Akbas, Ferhat, and Egemen Genc. "Do Mutual Fund Investors Overweight the Probability of Extreme Payoffs in the Return Distribution?." *Journal of Financial and Quantitative Analysis* 55, no. 1 (2020): 223-261.

Almazan, Andres, Keith Brown, Murray Carlson, and David Chapman. "Why constrain your mutual fund manager?." *Journal of Financial Economics* 73, no. 2 (2004): 289-321.

Amihud, Yakov, and Ruslan Goyenko. "Mutual fund's R2 as predictor of performance." *The Review of Financial Studies* 26, no. 3 (2013): 667-694.

Ang, Andrew, Sergiy Gorovyy, and Gregory van Inwegen. "Hedge fund leverage." *Journal of Financial Economics* 102, no. 1 (2011): 102-126.

Aragon, George. "Share restriction and asset pricing: Evidence from the hedge fund industry." *Journal of Financial Economics* 83, no. 1 (2007): 33-58.

Artavanis, Nikolaos T., Asli Eksi, and Gregory B. Kadlec. "Downside risk and mutual fund flows." (2019). Working paper.

Auer, Benjamin, and Frank Schuhmacher. "Liquid betting against beta in Dow Jones Industrial Average stocks." *Financial Analysts Journal* 71, no. 6 (2015): 30-43.

Baker, Malcolm, and Jeffrey Wurgler. “Investor sentiment and the cross-section of stock returns.” *The Journal of Finance* 61, no. 4 (2006): 1645-1680.

Baker, Malcolm, Brendan Bradley, and Jeffrey Wurgler. “Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly.” *Financial Analysts Journal* 67, no. 1 (2011): 40-54.

Bali, Turan G., Stephen J. Brown, and Mustafa O. Caglayan. "Do hedge funds' exposures to risk factors predict their future returns?." *Journal of Financial Economics* 101, no. 1 (2011): 36-68.

Bali, Turan G., Stephen J. Brown, and Mustafa O. Caglayan. "Macroeconomic risk and hedge fund returns." *Journal of Financial Economics* 114, no. 1 (2014): 1-19.

Bali, Turan G., Stephen J. Brown, and Mustafa O. Caglayan. "Upside potential of hedge funds as a predictor of future performance." *Journal of Banking & Finance* 98 (2019): 212-229.

Bali, Turan G., Stephen J. Brown, and Mustafa O. Caglayan. “Systematic risk and the cross section of hedge fund returns.” *Journal of Financial Economics* 106, no. 1 (2012): 114-131.

Bali, Turan G., Suleyman Gokcan, and Bing Liang. "Value at risk and the cross-section of hedge fund returns." *Journal of Banking & Finance* 31, no. 4 (2007): 1135-1166.

Bali, Turan, Stephen Brown, Scott Murray, and Yi Tang. “A lottery-demand-based explanation of the beta anomaly.” *Journal of Financial and Quantitative Analysis* 52, no. 6 (2017): 2369-2397.

Barber, Brad M., Terrance Odean, and Lu Zheng. "Out of sight, out of mind: The effects of expenses on mutual fund flows." *The Journal of Business* 78, no. 6 (2005): 2095-2120.

Barber, Brad M., Xing Huang, and Terrance Odean. "Which factors matter to investors? Evidence from mutual fund flows." *The Review of Financial Studies* 29, no. 10 (2016): 2600-2642.

Barth, Daniel, Juha Joenväärä, Mikko Kauppila, and Russ Wermers. "The Hedge Fund Industry is Bigger (and has Performed Better) Than You Think." (2020). Working paper.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song. "What do mutual fund investors really care about?." (2019). Working paper.

Berk, Jonathan, and Jules van Binsbergen. “Assessing asset pricing models using revealed preference.” *Journal of Financial Economics* 119, no. 1 (2016): 1-23.

Berk, Jonathan, and Jules van Binsbergen. “Measuring skill in the mutual fund industry.” *Journal of Financial Economics* 118, no. 1 (2015): 1-20.

Bhardwaj, Geetesh, Gary Gorton, and Geert Rouwenhorst. “Fooling some of the people all of the time: The inefficient performance and persistence of commodity trading advisors.” *Review of Financial Studies* 27, no. 11 (2014): 3099-3132.

Black, Fischer, Michael Jensen, and Myron Scholes. “The capital asset pricing model: Some empirical tests.” In Studies in the Theory of Capital Markets, Praeger, New York, (1972).

Blackburn, Douglas W., William N. Goetzmann, and Andrey D. Ukhov. "Risk aversion and clientele effects". (2009). Working paper.

Blitz, David. “Are hedge funds on the other side of the low-volatility trade?” *The Journal of Alternative Investments* 21, no. 1 (2018): 17-26.

Bodnaruk, Andriy, Bekhan Chokaev, and Andrei Simonov. "Downside risk timing by mutual funds." *The Review of Asset Pricing Studies* 9, no. 1 (2019): 171-196.

Boguth, Oliver, and Mikhail Simutin. “Leverage constraints and asset prices: Insights from mutual fund risk taking.” *Journal of Financial Economics* 127, no. 2 (2018): 323-341.

Bollen, Nicolas P.B., Juha Joenväärä, and Mikko Kauppila. “Hedge Fund Performance: End of an Era?” (2020). Working paper.

Brunnermeier, Markus, and Stefan Nagel. “Hedge funds and the technology bubble.” *The Journal of Finance* 59, no. 5 (2004): 2013-2040.

Caglayan, Mustafa, Umut Celiker, and Gokhan Sonaer. “Hedge fund vs. non-hedge fund institutional demand and the book-to-market effect.” *Journal of Banking & Finance* 92, (2018): 51-66.

Cao, Charles, Bing Liang, Andrew Lo, and Lubomir Petrasek. “Hedge fund holdings and stock market efficiency.” *The Review of Asset Pricing Studies* 8, no. 1 (2017): 77-116.

Cao, Charles, Yong Chen, Bing Liang, and Andrew W. Lo. "Can hedge funds time market liquidity?." *Journal of Financial Economics* 109, no. 2 (2013): 493-516.

Cao, Charles, Yong Chen, William Goetzmann, and Bing Liang. “Hedge funds and stock price formation.” *Financial Analysts Journal* 74, no. 3 (2018): 54-68.

Carhart, Mark. “On persistence in mutual fund performance.” *The Journal of Finance* 52, no. 1 (1997): 57-82.

Cederburg, Scott, and Michael O’Doherty. “Does it pay to bet against beta? On the conditional performance of the beta anomaly.” *The Journal of Finance* 71, no. 2 (2016): 737-774.

Chen, Yong, and Bing Liang. “Do market timing hedge funds time the market?.” *Journal of Financial and Quantitative Analysis* 42, no. 4 (2007): 827-856.

Choi, Jaewon, and Mathias Kronlund. “Reaching for Yield in Corporate Bond Mutual Funds.” *The Review of Financial Studies* 31, no. 5 (2018): 1930-1965.

Christoffersen, Susan, and Mikhail Simutin. “On the demand for high-beta stocks: Evidence from mutual funds.” *The Review of Financial Studies* 30, no. 8 (2017): 2596-2620.

Clifford, Christopher, Bradford Jordan, and Timothy Riley. “Do absolute-return mutual funds have absolute returns?.” *The Journal of Investing* 22, no. 4 (2013): 23-40.

Cochrane, John. “Presidential address: Discount rates.” *The Journal of Finance* 66, no. 4 (2011): 1047-1108.

Cooper, Michael J., Huseyin Gulen, and P. Raghavendra Rau. "Changing names with style: Mutual fund name changes and their effects on fund flows." *The Journal of Finance* 60, no. 6 (2005): 2825-2858.

Cremers, KJ Martijn, and Antti Petajisto. "How active is your fund manager? A new measure that predicts performance." *The Review of Financial Studies* 22, no. 9 (2009): 3329-3365.

Cremers, KJ Martijn, Jon A. Fulkerson, and Timothy B. Riley. "Challenging the conventional wisdom on active management: A review of the past 20 years of academic literature on actively managed mutual funds." *Financial Analysts Journal* 75, no. 4 (2019): 8-35.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz. "Should benchmark indices have alpha? Revisiting performance evaluation." *Critical Finance Review*, no. 2 (2012):1-48.

Cremers, Martijn. "Active share and the three pillars of active management: skill, conviction, and opportunity." *Financial Analysts Journal* 73, no. 2 (2017): 61-79.

Culkin, Robert, and Sanjiv R. Das. "Machine learning in finance: the case of deep learning for option pricing." *Journal of Investment Management* 15, no. 4 (2017): 92-100.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers. "Measuring mutual fund performance with characteristic-based benchmarks." *The Journal of Finance* 52, no. 3 (1997): 1035-1058.

DeWoody, Madeline. “A tale of two benchmarks: Performance benchmarking in the mutual fund industry.” (2019). Working paper.

Diebold, F. Mariano, and Roberto Mariano. "R.(1995). Comparing predictive accuracy." *Journal of Business & economic statistics* 20, no. 1 (1995).

Duanmu, Jun, Alexey Malakhov, and William McCumber. “Beta active hedge fund management.” *Journal of Financial and Quantitative Analysis* 53, no. 6 (2018): 2525-2558.

Edelman, Daniel, William Fung, David A. Hsieh, and Narayan Y. Naik. "Funds of hedge funds: performance, risk and capital formation 2005 to 2010." *Financial Markets and Portfolio Management* 26, no. 1 (2012): 87-108.

Edelman, Daniel, William Fung, David Hsieh, and Narayan Naik. “Fund of hedge funds: Performance, risk, and capital formation.” *Financial Markets and Portfolio Management* 26, no. 1 (2012): 87-108.

Efron, Bradley, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. “Least angle regression.” *The Annals of Statistics* 32, no. 2 (2004): 407-499.

Evans, Richard B. "Mutual fund incubation." *The Journal of Finance* 65, no. 4 (2010): 1581-1611.

Evans, Richard B., and Yang Sun. "Models or stars: The role of asset pricing models and heuristics in investor risk adjustment." *The Review of Financial Studies* (2020): 0893-9454.

Fama, Eugene F., and Kenneth R. French. "Luck versus skill in the cross-section of mutual fund returns." *The Journal of Finance* 65, no. 5 (2010): 1915-1947.

Fama, Eugene F., and R. Kenneth. "French, 1993, Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 33, no. 1 (1993): 3-56.

Ferson, Wayne, and Rudi Schadt. “Measuring fund strategy and performance in changing economic conditions.” *The Journal of Finance* 51, no. 2 (1996): 425-461.

Frazzini, Andrea, and Lasse Pedersen. “Betting against beta.” *Journal of Financial Economics* 111, no. 1 (2014): 1-25.

Frazzini, Andrea, David Kabiller, and Lasse Pedersen. “Buffett’s alpha.” *Financial Analysts Journal* 74, no. 4 (2018): 35-55.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. “The elements of statistical learning”. Vol. 1, no. 10. New York: *Springer series in statistics*, 2001.

Friend, Irwin, and Marshall Blume. “Measurement of portfolio performance under uncertainty.” *The American Economic Review* 60, no. 4 (1970): 561-575.

Fung, William, and David A. Hsieh. "Empirical characteristics of dynamic trading strategies: The case of hedge funds." *The Review of Financial Studies* 10, no. 2 (1997): 275-302.

Fung, William, and David A. Hsieh. "Hedge fund benchmarks: A risk-based approach." *Financial Analysts Journal* 60, no. 5 (2004): 65-80.

Fung, William, and David A. Hsieh. "The risk in hedge fund strategies: Theory and evidence from trend followers." *The Review of Financial Studies* 14, no. 2 (2001): 313-341.

Fung, William, David A.Hsieh, Narayan Naik, and Tarun Ramadorai. “Hedge funds: Performance, risk, and capital formation.” *The Journal of Finance* 63, no. 4 (2008): 1777-1803.

Goetzmann, William, Jonathan Ingersoll, and Zoran Ivkovic. “Monthly measurement of daily timers.” *Journal of Financial and Quantitative Analysis* 35, no. 3 (2000)” 257-290.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. "Empirical asset pricing via machine learning." *The Review of Financial Studies* 33, no. 5 (2020): 2223-2273.

Han, Xing. “Understanding the Performance of Components in Betting Against Beta.” *Critical Finance Review*, (2019): forthcoming.

Harvey, Campbell R., Yan Liu, and Heqing Zhu. "… and the cross-section of expected returns." *The Review of Financial Studies* 29, no. 1 (2016): 5-68.

Haugen, Robert, and A. James Heins. “On the evidence supporting the existence of risk premiums in the capital market.” (1972). Working paper.

Haugen, Robert, and A. James Heins. “Risk and the rate of return on financial assets: Some old wine in new bottles.” *Journal of Financial and Quantitative Analysis* 10, no. 5 (1975): 775-784.

Henriksson, Roy, and Robert Merton. “On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills.” *Journal of Business* (1981): 513-533.

Hong, Harrison, and David Sraer. “Speculative betas.” *The Journal of Finance* 71, no. 5 (2016): 2095-2144.

Hou, Kewei, Chen Xue, and Lu Zhang. “Replicating anomalies.” *The Review of Financial Studies* 33, no. 5 (2020): 2019-2133.

Huang, Jennifer C., Kelsey D. Wei, and Hong Yan. "Investor learning and mutual fund flows.". (2012). Working paper.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang. "Risk shifting and mutual fund performance." *The Review of Financial Studies* 24, no. 8 (2011): 2575-2616.

Huang, Jing-Zhi, and Ying Wang. “Should investors invest in hedge fund-like mutual funds? Evidence from the 2007 financial crisis.” *Journal of Financial Intermediation* 22, no. 3 (2013): 482-512.

Ibbotson, Roger, Peng Chen, and Kevin Zhu. “The ABCs of hedge funds: Alphas, betas, and costs.” *Financial Analysts Journal* 67, no. 1 (2011): 15-25.

Jagannathan, Ravi, Alexey Malakhov, and Dmitry Novikov. “Do hot hands exist among hedge fund managers? An empirical evaluation.” *The Journal of Finance* 65, no. 1 (2010): 217-255.

Jain, Prem C., and Joanna Shuang Wu. "Truth in mutual fund advertising: Evidence on future performance and fund flows." *The Journal of Finance* 55, no. 2 (2000): 937-958.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. “An introduction to statistical learning.” Vol. 112. New York: springer, 2013.

Joenväärä, Juha and Kauppila, Mikko and Kosowski, Robert and Tolonen, Pekka, “Hedge Fund Performance: Are Stylized Facts Sensitive to Which Database One Uses?” *Critical Finance Review* (2020), forthcoming

Joenvaara, Juha, Robert Kosowski, and Pekka Tolonen. “The effect of investment constraints on hedge fund investor returns.” *Journal of Financial and Quantitative Analysis* 54, no. 4 (2019): 1539-1571.

Jones, Christopher S., and Haitao Mo. "Out-of-sample performance of mutual fund predictors." *The Review of Financial Studies* (2016).

Jordan, Bradford D., and Timothy B. Riley. "Volatility and mutual fund manager skill." *Journal of Financial* Economics 118, no. 2 (2015): 289-298.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. "On the industry concentration of actively managed equity mutual funds." *The Journal of Finance* 60, no. 4 (2005): 1983-2011.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp. "Time-varying fund manager skill." *The Journal of Finance* 69, no. 4 (2014): 1455-1484.

Kahneman, Daniel, and Amos Tversky. "Prospect theory: An analysis of decision under risk." *Econometrica* 47, no. 2 (1979): 263-292.

Kaniel, Ron, and Robert Parham. "WSJ Category Kings–The impact of media attention on consumer and mutual fund investment decisions." *Journal of Financial Economics* 123, no. 2 (2017): 337-356.

Karagiannis, Nikolaos, and Konstantinos Tolikas. "Tail risk and the cross-section of mutual fund expected returns." *Journal of Financial and Quantitative Analysis* 54, no. 1 (2019): 425-447.

Karceski, Jason. “Return-chasing behavior, mutual funds, and beta’s death.” *Journal of Financial and Quantitative Analysis* 37, no. 4 (2002): 559-594.

Khandani, Amir E., Adlar J. Kim, and Andrew W. Lo. "Consumer credit-risk models via machine-learning algorithms." *Journal of Banking & Finance* 34, no. 11 (2010): 2767-2787.

Kokkonen, Joni, and Matti Suominen. “Hedge funds and stock market efficiency.” *Management Science* 61, no. 12 (2015): 2890-2904.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White. "Can mutual fund “stars” really pick stocks? New evidence from a bootstrap analysis." *The Journal of Finance* 61, no. 6 (2006): 2551-2595.

Kosowski, Robert, Narayan Naik, and Melvyn Teo. “Do hedge funds deliver alpha? A Bayesian and bootstrap analysis.” J*ournal of Financial Economics* 84, no. 1 (2007): 229-264.

Lee, Jung Hoon, Charles Trzcinka, and Shyam Venkatesan. “Do Portfolio Manager Contracts Contract Portfolio Management?.” *Journal of Finance* 74, no. 5 (2019): 2543-2577.

Li, Yongjia, and Alexey Malakhov. “Bringing Order to Chaos: Capturing Relevant Information with Hedge Fund Factor Models.” (2017). Working paper.

Liu, Jianan, Robert Stambaugh, and Yu Yuan. “Absolving beta of volatility’s effects.” *Journal of Financial Economics* 128, no. 1 (2018): 1-15.

Ma, Linlin, and Yuehua Tang. "Portfolio manager ownership and mutual fund risk taking." *Management Science* 65, no. 12 (2019): 5518-5534.

Ma, Linlin, Yuehua Tang, and Juan-Pedro Gomez. “Portfolio manager compensation in the U.S. mutual fund industry.” *The Journal of Finance* 74, no. 2 (2019): 587-638.
Malkiel, Burton, and Atanu Saha. “Hedge funds: Risk and return.” *Financial Analysts Journal* 61, no. 6 (2005): 80-88.

Mamaysky, Harry, Matthew Spiegel, and Hong Zhang. "Estimating the dynamics of mutual fund alphas and betas." *The Review of Financial Studies* 21, no. 1 (2008): 233-264.

McLean, R. David, and Jeffrey Pontiff. "Does academic research destroy stock return predictability?." *The Journal of Finance* 71, no. 1 (2016): 5-32.

Novy-Marx, Robert, and Mihail Velikov. “A taxonomy of anomalies and their trading costs.” *Review of Financial Studies* 29, no. 1 (2016): 104-147.

Novy-Marx, Robert, and Mihail Velikov. “Betting against betting against beta.” (2019).Working paper.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor. "Do funds make more when they trade more?." *The Journal of Finance* 72, no. 4 (2017): 1483-1528.

Patton, Andrew J. "Are “market neutral” hedge funds really market neutral?." *The Review of Financial Studies* 22, no. 7 (2009): 2495-2530.

Patton, Andrew, and Tarun Ramadorai. “On the high-frequency dynamics of hedge fund risk exposures.” *The Journal of Finance* 68, no. 2 (2013): 597-635.

Phillips, Blake, Kuntara Pukthuanthong, and Raghavendra Rau. “Size does not matter: Diseconomies of scale in the mutual fund industry revisited.” *Journal of Banking & Finance* 88, (2018): 357-365.

Polkovnichenko, Valery, Kelsey D. Wei, and Feng Zhao. "Cautious risk takers: Investor preferences and demand for active management." *The Journal of Finance* 74, no. 2 (2019): 1025-1075.

Rapach, David E., Jack K. Strauss, and Guofu Zhou. "International stock return predictability: what is the role of the United States?." *The Journal of Finance* 68, no. 4 (2013): 1633-1662.

Riley, Timothy B. "Can mutual fund stars still pick stocks?: A replication and extension of Kosowski, Timmermann, Wermers, and White (2006)." *Critical Finance Review* (2019).

Sadka, Ronnie. "Liquidity risk and the cross-section of hedge-fund returns." *Journal of Financial Economics* 98, no. 1 (2010): 54-71.

Sensoy, Berk. “Performance evaluation and self-designated benchmark indexes in the mutual fund industry.” *Journal of Financial Economics* 92, no. 1 (2009): 25-39.

Shive, Sophie, and Hayong Yun. “Are mutual funds sitting ducks?.” *Journal of Financial Economics* 107, no. 1 (2013): 220-237.

Sias, Richard, Harry Turtle, and Blerina Zykaj. “Hedge fund crowds and mispricing.” *Management Science* 62, no. 3 (2015): 764-784.

Sirri, Erik R., and Peter Tufano. "Costly search and mutual fund flows." *The Journal of Finance* 53, no. 5 (1998): 1589-1622.

Solomon, David H., Eugene Soltes, and Denis Sosyura. "Winners in the spotlight: Media coverage of fund holdings as a driver of flows." *Journal of Financial Economics* 113, no. 1 (2014): 53-72.

Sun, Zheng, Ashley W. Wang, and Lu Zheng. "Only winners in tough times repeat: Hedge fund performance persistence over different market conditions." *Journal of Financial and Quantitative Analysis* 53, no. 5 (2018): 2199-2225.

Sun, Zheng, Ashley Wang, and Lu Zheng. "The road less traveled: Strategy distinctiveness and hedge fund performance." *The Review of Financial Studies* 25, no. 1 (2012): 96-143.

Tibshirani, Robert. “Regression shrinkage and selection via the lasso.” *Journal of the Royal Statistical Society* 58, no. 1 (1996): 267-288.

Titman, Sheridan, and Cristian Tiu. “Do the best hedge funds hedge?.” *The Review of Financial Studies* 24, no. 1 (2010): 123-168.

Von Reibnitz, Anna. "When opportunity knocks: Cross-sectional return dispersion and active fund performance." *Critical Finance Review* 6, no. 2 (2017): 303-356.

Wang, Albert Y., and Michael Young. "Terrorist attacks and investor risk preference: Evidence from mutual fund flows." *Journal of Financial Economics* 137, no. 2 (2020): 491-514.